(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 3 523 319 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
28.06.2023 Bulletin 2023/26

(21) Application number: **17790700.3**

(22) Date of filing: **10.10.2017**

(51) International Patent Classification (IPC):
C07K 14/415 *(2006.01)* ***C12N 15/82*** *(2006.01)*
C12Q 1/6895 *(2018.01)*

(52) Cooperative Patent Classification (CPC):
C07K 14/415; C12N 15/8286; C12Q 1/6895;
C12Q 2600/13

(86) International application number:
PCT/EP2017/075849

(87) International publication number:
WO 2018/069343 (19.04.2018 Gazette 2018/16)

(54) NUCLEIC ACID ENCODING SM1 RESISTANCE TO ORANGE WHEAT BLOSSOM MIDGE AND METHOD OF USE

NUKLEINSÄURE ZUR CODIERUNG DER SM1-RESISTENZ GEGEN DIE ORANGEFARBENE WEIZENGALLMÜCKE UND VERFAHREN ZUR VERWENDUNG

RÉSISTANCE SM1 DE CODAGE D'ACIDE NUCLÉIQUE À LA CÉCIDOMYIE ORANGÉE DU BLÉ ET SON PROCÉDÉ D'UTILISATION

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **10.10.2016 EP 16306334**

(43) Date of publication of application:
14.08.2019 Bulletin 2019/33

(73) Proprietor: **Limagrain Europe**
63360 Saint-Beauzire (FR)

(72) Inventors:
- **BERRY, Simon Thimothy**
  Litcham
  Norfolk PE32 2QQ (GB)
- **COMADRAN, Jordi**
  63360 Gerzat (FR)
- **SPECEL, Sébastien**
  63360 Gerzat (FR)

(74) Representative: **Regimbeau**
20, rue de Chazelles
75847 Paris Cedex 17 (FR)

(56) References cited:
- **DATABASE EMBL [Online] 26 February 2001 (2001-02-26), "WHE2063_C10_E19ZS Wheat salt-stressed sheath cDNA library Triticum aestivum cDNA clone WHE2063_C10_E19, mRNA sequence.", XP002766805, retrieved from EBI accession no. EM_EST:BG313458 Database accession no. BG313458**
- **KASSA MULUALEM T ET AL: "A saturated SNP linkage map for the orange wheat blossom midge resistance geneSm1", THEORETICAL AND APPLIED GENETICS ; INTERNATIONAL JOURNAL OF PLANT BREEDING RESEARCH, SPRINGER, BERLIN, DE, vol. 129, no. 8, 9 May 2016 (2016-05-09), pages 1507-1517, XP036003573, ISSN: 0040-5752, DOI: 10.1007/S00122-016-2720-4 [retrieved on 2016-05-09] cited in the application**
- **THOMAS J ET AL: "Chromosome location and markers of Sm1: a gene of wheat that conditions antibiotic resistance to orange wheat blossom midge", MOLECULAR BREEDING, KLUWER ACADEMIC PUBLISHERS, DO, vol. 15, no. 2, 1 February 2005 (2005-02-01), pages 183-192, XP019258713, ISSN: 1572-9788 cited in the application**

Processed by Luminess, 75001 PARIS (FR)

(Cont. next page)

- **POZNIAK C J ET AL: "CDC Carbide durum wheat", CANADIAN JOURNAL OF PLANT SCIENCE, vol. 95, no. 5, September 2015 (2015-09), pages 1007-1012, XP002766806,**
- **DATABASE EMBL [Online] 19 September 2011 (2011-09-19), "TSA: Triticum aestivum contig13527.TraeRec mRNA sequence.", retrieved from EBI accession no. EM_TSA:JP218692 Database accession no. JP218692**
- **DATABASE EMBL [Online] 19 August 2012 (2012-08-19), "TSA: Triticum aestivum Ta_Contig47067.ansp mRNA sequence.", retrieved from EBI accession no. EM_TSA:JV939981 Database accession no. JV939981**
- **DATABASE EMBL [Online] 28 July 2014 (2014-07-28), "Triticum aestivum chromosome 3B, genomic scaffold, cultivar Chinese Spring", retrieved from EBI accession no. EM_STD:HG670306 Database accession no. HG670306**
- **DATABASE EMBL [Online] 19 August 2012 (2012-08-19), "TSA: Triticum aestivum Ta_Contig8850.ansp mRNA sequence.", retrieved from EBI accession no. EM_TSA:JV871805 Database accession no. JV871805**

Remarks:

The complete document including Reference Table(s) and the Sequence Listing(s) can be downloaded from the EPO website

Description

BACKGROUND:

Economic importance of orange wheat blossom midge

[0001] Orange wheat blossom midge (OWBM; Sitodiplosis mosellana Géhin) is a serious, economic insect pest of both bread (Triticum aestivum) and durum wheats (Triticum durum) in all the main temperate cereal-growing regions of the world, including North America, Canada, Europe and Asia (Harris et al, 2003). Apart from T. aestivum, OWBM has been shown to oviposit and survive on accessions of sixteen other Triticum species (Wise et al, 2001), as well as rye, barley, and a number of wild grasses (Zheng, 1965). The feeding damage caused by the OWBM larvae not only results in significant yield losses through the abortion of flowers, grain deformation and reduced specific weight, but it also makes the grain more vulnerable to fungal infection (Oakley, 1994) and reduces quality by lowering the Hagberg Falling Number (Miller and Halton, 1960). In the United Kingdom, crop losses exceeded £30 million in 1993 (Oakley, 1994) and £60 million in 2004, in spite of spraying more than 500,000 hectares of wheat with insecticides (Oakley et al, 2005). Chavalle et al (2015) demonstrated that the relationship between yield and the number of larvae was logarithmic, suggesting that the reduction in yield was caused either by the damage inflicted by young larvae that died at the start of their development or by the activation of costly resistant reactions in the plant.

Life cycle of OWBM (Anon, 2016)

[0002] OWBM larvae over-winter in the soil as cocoons until diapause is broken (70 days of less than 10°C) and then the activated larvae move towards the soil surface. The larvae pupate once rainfall wets the soil to a depth of 10mm and the soil temperature rises above 13°C. The duration of the pupal stage varies according to temperature, but typically lasts between 2 to 4 weeks. Increasing temperatures, following rainfall, stimulate the hatch of the adults from pupae. Shortly after emergence, females emit a sexual pheromone that attracts the males, which has been identified as (2S,7S)-2,7-nonanediyl dibutyrate (Gries et al. 2000). After mating at the emergence site, the female midge fly off in search of a host plant. Wheat only becomes infested by OWBM if it is exposed to ovipositing females at the time the heads emerge from the flag leaves until anthesis - about 7 days (Elliot and Mann, 1996).

[0003] Females are active at dusk, preferring warm, calm, humid evenings, but they only live for 3 to 7 days. Eggs are laid singly or in small groups (normally 1 to 6 eggs) underneath the glumes, but females tend to lay eggs where others have already oviposited. Pre-anthesis heads receive nine times more eggs than post anthesis heads (Ding and Lamb, 1999) as the female midge are attracted by volatile compounds released from pre-anthesis wheat spikes (Birkett et al. 2004). Eggs hatch after 4 to 7 days and the larvae crawl into the florets and feed on the surface of the developing kernels. The larvae will continue feeding for about 2 to 3 weeks and then, following heavy rainfall, they drop out of the spikes back into the soil to form new cocoons, which can remain viable in the ground for 10 years or more.

Chemical control of OWBM

[0004] Control of OWBM in wheat crops can be achieved by spraying chlorpyrifos (IUPAC name: O,O-diethyl O-3,5,6-trichloropyridin-2-yl phosphorothioate), which is an organophosphorus pesticide acting on the nervous system of insects via the inhibition of acetylcholinesterase. Unfortunately this insecticide only kills the adult midge and eggs, but not the larvae within the wheat florets and so the timing of the spray application is critical. The development of pheromone traps, using (2S,7S)-2,7-nonanediyl dibutyrate, (Bruce et al. 2007) has enabled farmers to readily assess the level of midge infestation within their fields. Midge counts can be used as part of a grower's risk assessment to determine whether or not to spray (http://uk.dowagro.com/wp-content/uploads/2014/05/Wheat-Blossom-Midge-Risk-Assessment-Chart.pdf); however the use of chlorpyrifos has now been banned in the UK, as of April the 1ST 2016 (http://cereals.ahdb.org.uk/chlorpyrifos). The EU has also recently lowered the maximum residue levels for chlorpyrifos on all food crops (Commission Regulation (EU) 2016/60).

Biological control of OWBM

[0005] There are a number of natural predators of OWBM, but data is limited on their effectiveness (Doane et al, 2013), apart from the parasitic wasp Macroglenes penetrans, which plays an important role in reducing OWBM infestations in Western Canada (Olfert et al, 2009).

Plant resistance mechanisms to insect pests

[0006] Many different mechanisms for increased resistance to insect herbivory have been identified in plants (War et al, 2012), but these generally fall into three categories:

1) Tolerance - a polygenic trait enabling the plant to withstand or recover from insect damage, but it does not adversely affect the growth or survival of the insect

2) Antixenosis - a non-preference reaction of insects to a resistant plant caused by plant morphology (e.g. trichomes) or chemical factors that adversely affect insect behavior, often resulting in delayed acceptance or outright rejection.

3) Antibiosis (90% of all described insect resistances) - a resistant plant adversely affects the life cycle of the insect either through its survival rate, development or fecundity.

[0007] Many insect resistance QTL have been genetically mapped in different crop plant species (Smith and Clement, 2012), but to date, only a handful of the underlying genes have been successfully cloned. These include: Mi-1.2 from tomato (Rossi et al, 1998), Bph14 from rice (Du et al, 2009) and Vat from melon (Dogimont et al. 2014), all of which have been identified as coiled coil-nucleotide binding site-leucine rich repeats (CC-NBS-LRRs), also known as CNL or non-TIR NBS-LRR's (Meyers et al, 1999).

Genetic resistance to OWBM

Antixenosis:

[0008] For the female wheat midge, there are at least three points, in its behavioral sequence leading to oviposition, that are subject to deterrence: 1) locating and landing on the plant, 2) accepting the plant and initiating probing and 3) finally initiating the oviposition event itself (Gharalari et al, 2011). Oviposition deterrence has been demonstrated by Lamb et al, (2001, 2002, 2003), but its multi-genic nature and environmentally influenced expression make it difficult to incorporate this type of trait into breeding programs (Gharalari et al. 2009a). It is thought that the female midge uses fine-scale features and surface chemicals, including volatiles, to select oviposition sites on the wheat spike (Gharalari et al. 2011). Volatile compounds produced by deterrent genotypes, or at less preferred growth stages (e.g. post-anthesis), do reduce oviposition on preferred genotypes and may cause the females to lay their eggs further from potential larval feeding sites (Lamb et al. 2003; Gharalari et al, 2011). There are no known morphological traits that account for oviposition preferences, apart from a small effect of inter-spikelet distance (Gharalari et al. 2009b).

Antibiosis:

[0009] Wheat defends itself in at least two ways against the feeding larvae, both of which appear to be associated with increased production of phenolic acids in the developing seed coat (Ding et al, 2000):

1) The feeding site is suitable for larvae only at an early stage of seed development (Ding and Lamb, 1999). Ten days post-anthesis all seeds become antibiotic to newly hatched larvae. Consequently wheat midge larvae feed on susceptible, young seeds for 10 to 12 days beginning about the time that pollination occurs (Ding and Lamb, 1999).

2) Some wheat genotypes appear to exhibit a hypersensitive reaction on the seed surface that kills the feeding larvae (Lamb et al. 2000) leaving so-called "wound-plugs". This hypersensitive reaction is rapidly induced by larval feeding (Ding et al. 2000) and decreases the survival of the first larval instar by 99% (Lamb et al. 2000). McKenzie et al (2002) demonstrated that this hypersensitive response was controlled by a single, partially dominant gene, which the authors named Sm1. To date, very few alternative genetic resistances to Sm1 have been genetically mapped. One example, published by Blake et al, (2011), identified a QTL (QSm.mst-1A) in the spring wheat variety Reeder, which reduced OWBM infestation by 42% in near isogenic lines with and without the QTL region.

Sm1 antibiotic resistance locus:

[0010] Lamb et al (2015) have shown that the expression of the Sm1 locus can vary between wheat varieties and the authors suggest that this may be related to the rate of production and/or the types of phenolic acids produced at the seed surface once resistance has been induced (Ding et al. 2000). Lamb et al (2015) also noted that the inheritance of Sm1 was independent of other resistance mechanisms such as oviposition deterrence.

[0011] Thomas et al (2005) mapped the Sm1 locus to the short-arm of chromosome 2B in wheat and developed a linked, dominant SCAR marker (WM1) for use in marker-assisted selection (MAS). More recently, Kassa et al (2016) have refined the map position of the Sm1 locus and discovered that this genomic region in wheat is inverted when to compared to the rice and Brachypodium genomes. They also speculated that Sm1 might be a traditional resistance gene of the CC-NBS-LRR type, but they were only able identify one candidate resistance gene of the NB-ARC type in B. distachyon and none were found in the syntenic region in rice. Kassa et al (2016) also identified two SNP markers (kwm707 and kwm693) that could be used to predict the presence of Sm1 locus in diverse panels of wheat varieties, suggesting a single origin for Sm1; however these markers were not 100% diagnostic.

[0012] Therefore there is an urgent need to identify perfect markers for Sm1 via the cloning and sequencing of this locus.

BRIEF DESCRIPTION OF THE DRAWINGS:

[0013]

Figure 1: RGA 1 and RGA 2 representation showing the relative positions of the functional domains identified with InterProScan: (1) Coiled coil; (2) NBS-ARC; (3) LRR; (4) NAM; (5) pKinase; (6) PapD-like (also called Major Sperm Protein (MSP) domain).

Figure 2. The physical alignment between BAC 715D09 and the IWGSC WGA scaffolds: IWGSCWGAV02_2BS_scaffold14096 and IWGSCWGAV02_2BS_ scaffold11627. RGA.CS refers to the Chinese Spring homologs of the RGA genes annotated in the BAC 715D09.

Figure 3. An Integrative Genomics Viewer (IVG) screenshot showing the alignment of genomic (gDNA-seq) and transcriptomic (mRNA-seq) Illumina reads to the region of BAC clone 715D09 containing RGA 1.

DETAILED DESCRIPTION OF THE INVENTION:

[0014] The present invention is as defined in the appended set of claims.

[0015] The applicant has identified the genetic determinism driving OWBM resistance and is providing different means to improve or newly confer resistance to OWBM in plants. In particular the applicant is providing the nucleic acid encoding for protein correlated with the resistance of the Sm1 locus in wheat and uses of these nucleic acids.

[0016] The invention also relates to constructs which can be used as a transgene for obtaining transgenic plants that have improve or newly confer resistance with regards to isogenic plants that do not contain said transgene. The invention also relates to transgenic plants containing such constructs and the method for obtaining such plants.

[0017] Finally, markers and method for identifying said locus in plants are also provided.

[0018] In the context of the present invention a plant shall mean a monocotyledon and more preferentially a cereal, in particular from Triticum species (wheat), maize, rice, barley, sorghum, millet, oats, rye.

[0019] Wheat is the preferred cereal according to the invention.

[0020] As intended herein, two plants are said to be "isogenic" with regards to a transgene when they differ at very few loci (less than 20, more preferably less than 10), and when one does carry the transgene, while the other does not.

[0021] One embodiment of the invention is the nucleic acid encoding for a protein conferring orange wheat blossom midge resistance, wherein the nucleic acid comprises at least one of the group consisting of:

a) Group 1:

i. A nucleic acid comprising SEQ ID NO:4 or 5, or
ii. A nucleic acid encoding an amino acid sequence comprising SEQ ID NO:6, or
iii. A nucleic acid encoding an amino-acid sequence comprising at least one CC motif, one NBS-ARC motif and one LRR motif and wherein the sequence encoding the LRR motif is having at least 82% identity with the nucleic acid fragment as depicted in SEQ ID NO:8, and

b) Group 2:

i. A nucleic acid comprising SEQ ID NO:1 or 2, or
ii. A nucleic acid encoding an amino acid sequence comprising SEQ ID NO:3, or
iii. A nucleic acid encoding an amino-acid sequence comprising at least one CC motif, one NBS-ARC motif and one LRR motif and wherein the sequence encoding the LRR motif is having at least 78% identity with the nucleic acid fragment as depicted in SEQ ID NO:7,

[0022] The nucleic acids encoding the Resistance Gene Analogs RGA 1 and RGA 2 are both perfectly correlated to OWBM resistance in plants and more specifically in wheat. Each nucleic acid could be used alone or in combination for improving or conferring a new resistance in a plant.

[0023] As disclosed herein, SEQ ID NO:3 provides a wheat RGA 1 protein, SEQ ID NO:1 provides the wheat genomic sequence encoding for RGA 1 and SEQ ID NO:2 provides the wheat coding sequence (from a spliced RNA or a cDNA) encoding for RGA 1.

[0024] Similarly, SEQ ID NO:6 provides a wheat RGA 2 protein, SEQ ID NO:4 provides a wheat genomic sequence encoding for RGA 2 and SEQ ID NO:5 provides a wheat coding sequence encoding for RGA 2.

[0025] The invention also relates to the protein encoded by the nucleic acid of the invention comprising the group 1 and/or the group 2 as defined above. In a preferred embodiment, the invention relates to the sequences SEQ ID NO: 3 and SEQ ID NO: 6.

[0026] Both RGA proteins are comprising CC (coiled-coil), nucleotide-binding site (NBS-ARC) and leucine-rich repeat (LRR) domains, as well as variable amino- and carboxy-terminal domain (Sekhwal et al, 2015). Figure 1 provides the relative position of these different domains onto RGA 1 and RGA 2 proteins.
Interestingly, RGA 2 protein also comprises a NAM domain, a pkinase domain and a PadD-like domain which could be involved in defense systems in plants.

[0027] Different RGAs have been characterized by genome-wide identification and genetic mapping. Hundreds of NBS-LRR containing RGAs have been identified in different plants like Arabidopsis, barley, rice, maize and sorghum.

[0028] For the characterization of these proteins, many different software packages are publicly available to predict the position of the different motifs (i.e. CC, NBS, LRR and kinase domain) in a polypeptide such as, but not limited to, InterProScan (https://www.ebi.ac.uk/interpro/search/sequence-search), MEME (meme-suite-.org, Bailey and Elkan (1994)) or pfam_scan.pl (bit.ly/1M41KRu). Common software used for RGA domain and motif identification are also listed in Sekhwal et al, (2015).

[0029] More specifically, the LRR motif of RGA 1 SEQ ID NO: 3 is corresponding to the fragment defined by the amino acid position 599 to the amino acid position 989 while SEQ ID NO:3 is 1435 amino acids long.

[0030] The LRR motif of RGA 2 SEQ ID NO: 6 is corresponding to the fragment defined by the amino acid position 1310 to the amino acid position 1615 while SEQ ID NO:6 is 2303 amino acids long.

[0031] The invention also encompasses variant nucleic acid sequences encoding for variant proteins of RGA 1 or RGA 2 having one or more deletion/addition or one or more substitution compared to respectively SEQ ID NO:3 or SEQ ID NO: 6 and still possessing the function of improving or conferring a new OWBM resistance. Variant protein of RGA 1 will comprise at least one CC, one NBS-ARC, and one LRR motifs that are encoded by a nucleic acid having at least 78, 79, 80, 85, 90, 95, 96, 97, 98 or 99% identity with the SEQ ID NO:7.

[0032] Variant protein of RGA 2 will comprise at least one CC, one NBS-ARC and one LRR- motifs and are encoded by a nucleic acid having at least 82, 83, 84, 85, 90, 95, 96, 97, 98 or 99% identity with SEQ ID NO:8. Variant protein of RGA 2 are further comprising a kinase domain.

[0033] Variant nucleic acids is intended to mean natural variants identified in different cultivars or orthologous sequences identified in different species. Variants can also correspond to modifications introduced by mutagenesis in the nucleic acid sequence. These modifications can be made randomly or at specific sites. Different methods are well known to achieve such modification. One method can be based on random mutagenesis such as TILLING (Till et al, 2003), DNA shuffling (Stemmer, 1994) or on a targeted sequence modification provided by double-strand break technologies such as, but not limited to, TALENs (WO2011072246) or CRISPR cas9 (WO2013181440).

[0034] The variant nucleic acid can encode a variant protein if the modification at the nucleic acid creates a new codon leading to a new amino acid. It is then possible to screen for modification leading to a protein with altered conformation that promotes a more effective pathogen resistance.

[0035] Preferably, using these genome editing tools, it is expected to generate mutants by achieving whole domain modifications, for example via the "domain swapping" method, for further improving the resistance function.

[0036] The nucleic acid sequences may be identified from databases, by applying the BLASTN program (especially the BLASTN 2.2.30 program; Altschul et al, 1997; Altschul et al, 2005) preferably to SEQ ID NO: 1, 2, 4 or 5 using the following algorithm parameters:

Expected threshold: 10
Word size: 11
Max matches in a query range: 0
Gap Costs: Existence 5, Extension 2.
No filter for low complexity regions

[0037] Fragments of the nucleic acid are also disclosed herein but not as part of the present invention. Nucleic acid that are fragments comprises at least 20, 50, 100, 200, 300, 500 to 1,000 bp.

[0038] The nucleic acid can then be used in a construct under an operably linked heterologous promoter, which is also an aspect of the present invention.

[0039] As used herein, heterologous promoter means a promoter which does not originate from the same species from which the nucleic acid was derived, or the promoter is from the same species from which the nucleic acid was derived but has been modified to obtain a sequence different from the native sequence.

[0040] Operably linked means that there is a functional linkage between the regulatory element (the promoter) and the nucleic acid to allow the expression of the nucleic acid. Both elements can be separated by sequence that can enhance the expression of the nucleic acid like introns.

[0041] In a preferred embodiment, in the nucleic acid construct of the invention, the nucleic acid is cloned downstream of a heterologous promoter functional in a plant cell.

[0042] A promoter "active in plants" is a promoter that is able to drive expression of a gene operably linked thereto in a plant cell.

[0043] For being expressed, a sequence coding for RGA 1 or RGA 2 may be present under the control of a constitutive, tissue specific, developmentally regulated, inducible or meiosis promoter.

[0044] Although some promoters may have the same pattern of regulation when there are used in different species, it is often preferable to use monocotyledonous promoters in monocotyledons and dicotyledonous promoters in dicotyledonous plants.

[0045] In a preferred embodiment, said construct is under the control of a constitutive promoter.

[0046] Examples of constitutive promoters useful for expression include the 35S promoter or the 19S promoter (Kay et al, 1987), the rice actin promoter (McElroy et al, 1990), the pCRV promoter (Depigny-This et al, 1992), the CVMV promoter (Verdaquer et al. 1996), the ubiquitin 1 promoter of maize (Christensen and Quail, 1996), the regulatory sequences of the T-DNA of Agrobacterium tumefaciens, including mannopine synthase, nopaline synthase, octopine synthase.

[0047] More preferably the promoters used in the invention are those expressed during seed development such as the HMWG promoter (High Molecular Weight Glutenin) of wheat (Anderson and Greene 1989; Roberts et al., 1989), the waxy, zein or bronze promoters of maize, or the promoters disclosed in US 20150007360, US 20120011621, US 20100306876, US 20090307795 or US 20070028327.

[0048] Other suitable promoters could be used. It could be an inducible promoter, a developmentally regulated promoted or a tissue-specific promoter such as a leaf-specific promoter, a seed-specific, a BETL specific promoter and the like. Numerous tissue-specific promoters are described in the literature and any one of them can be used. One can cite the promoters disclosed in US 20130024998.

[0049] The invention also encompasses a vector containing the nucleic acid construct of the invention.

[0050] A vector, such as a plasmid, can thus be used for transforming host cells. The construction of vectors for transformation of host cells is within the capability of one skilled in the art following standard techniques.

[0051] The decision as to whether to use a vector for transforming a cell, or which vector to use, is guided by the method of transformation selected, and by the host cell selected.

[0052] Where a naked nucleic acid introduction method is used, then the vector can be the minimal nucleic acid necessary to confer the desired phenotype, without the need for additional sequences.

[0053] Possible vectors include the Ti plasmid vectors, shuttle vectors designed merely to maximally yield high numbers of copies, episomal vectors containing minimal sequences necessary for ultimate replication once transformation has occurred, transposon vectors, including the possibility of RNA forms of the gene sequences. The selection of vectors and methods to construct them are commonly known to persons of ordinary skill in the art and are described in general technical references (Mullis and Faloona, 1987).

[0054] For other transformation methods requiring a vector, the selection of an appropriate vector is relatively simple, as the constraints are minimal. The apparent minimal traits of the vector are that the desired nucleic acid be introduced in a relatively intact state. Thus, any vector which produces a plant carrying the introduced DNA sequence should be sufficient. Also, any vector which introduces a substantially intact RNA which can ultimately be converted into a stably maintained DNA sequence should be acceptable.

[0055] For transformation methods within a plant cell, one can cite methods of direct transfer of genes such as direct micro-injection into plant embryos, vacuum infiltration or electroporation, direct precipitation by means of PEG or the bombardment by gun of particles covered with the plasmid DNA of interest.

[0056] It is preferred to transform the plant cell with a bacterial strain, in particular Agrobacterium, in particular Agrobacterium tumefaciens. In particular, it is possible to use the method described by Ishida et al, (1996) for the transformation of Monocotyledons.

[0057] However, any additional attached vector sequences which confer resistance to degradation of the nucleic acid fragment to be introduced, which assists in the process of genomic integration or provides a means to easily select for those cells or plants which are actually, in fact, transformed are advantageous and greatly decrease the difficulty of selecting useable transgenic plants.

[0058] The vector can exist, for example, in the form of a phage, a plasmid or a cosmid. The construction of such expression vectors for transformation is well known in the art and uses standard techniques. Mention may be made of the methods described by Sambrook et al. (1989).

[0059] For transforming bacteria, a vector is generally defined as being a nucleic acid molecule that possesses elements that allows it to be maintained within said host cell (such as an origin of replication that works in this bacterial host cell).

[0060] The invention also encompasses a host cell containing at least the nucleic acid construct of the invention comprising group 1 and/or group 2 as described above.

[0061] The decision as to whether to use a given host cell, or which host cell to use, is guided by the method of transformation.

[0062] The host cell can be any prokaryotic or eukaryotic cell. Any of a large number of available and well-known host cells may be used in the practice of this invention. The selection of a particular host is dependent upon a number of factors recognized by the art. These include, for example, compatibility with the chosen expression vector, bio-safety and costs. Useful hosts include bacteria such as E. coli sp. or Agrobacterium. A plant host cell, may be also used, notably a Monocotyledon plant cell, particularly a cereal plant cell, selected in particular from the group consisting of maize, wheat, barley rice, and preferentially wheat.

[0063] More particularly, the host cell used in carrying out the invention is Agrobacterium tumefaciens, or Agrobacterium rhizogenes, according to the method described by Jouanin et al, (1987).

[0064] In a specific embodiment, said nucleic acid construct is stably integrated within the genome of said host cell. This embodiment is particularly interesting for plant host cells. Stable integration within the genome means that the expression cassette can be transmitted to the progeny of said host cell upon division.

[0065] The invention also encompasses a transgenic plant containing at least one cell containing the nucleic acid construct as defined above, preferably stably integrated within its genome.

[0066] A part of such a transgenic plant, in particular fruit, seed, grain or pollen, comprising such a cell or generated from such a cell, is also encompassed by the invention.

[0067] It is reminded that a whole plant can be regenerated from a single transformed plant cell. Thus, in a further aspect the present invention provides transgenic plants, or parts of them, including the expression cassette according to the invention. The regeneration can proceed by known methods.

[0068] The seeds which grow by fertilization from this plant, also contain this transgene in their genome.

[0069] Said plant or part of a plant according to the invention can be a plant or a part of it from various species, notably a Monocotyledons. Said plant is preferably a cereal plant.

[0070] Said plant is preferably selected from the group consisting of maize, rice, wheat and barley. In a preferred embodiment, said plant is wheat.

[0071] The invention thus relates in particular to a transgenic wheat, containing at least one cell comprising, stably integrated in its genome, the nucleic acid construct of the invention.

[0072] In a specific embodiment, said plant, in particular said wheat, comprises multiple cells containing, stably integrated in their genome, the nucleic acid construct of the invention. In this embodiment, it is possible that some cells of said plant do not contain the transgene.

[0073] In another embodiment, said transgene comprising the nucleic acid construct of the invention is present in all cells of said plant, in particular said wheat.

[0074] In another embodiment, the transgene is introduced within the plant cells such as being expressed transiently, or through a genetic construct not integrated in the genome. Thus, agro-infiltration or any other methods, such as injection or spray, are contemplated for transient expression.

[0075] Hybrid plants obtained by crossing plants according to the invention also form part of the invention, when they contain at least one cell containing the expression cassette of the invention.

[0076] Any plant as described above can contain one or more transgenes in addition to the cassette according to the invention. One may mention transgenes conferring male sterility, male fertility, resistance to a herbicide (notably glyphosate, glufosinate, imidazolinone, sulfonylurea, L-phosphinotricine, triazine, benzonitrile), resistance to insects (notably a transgene coding for a Bacillus thuringiensis toxin), tolerance to water stress. These plants can be obtained by crossing said plants of the invention with other plants containing said transgenes. Alternatively, plants can be co-transformed with an expression cassette containing several different transgenes, including the transgene of the invention.

[0077] Transgenic plants comprising an expression cassette according to the invention present an increased or improved resistance to OWBM as compared to control plants corresponding to non-transgenic plants not comprising said expression cassette.

[0078] In the present invention resistance to OWBM is measured as follows: plants are grown under field conditions and phenotyped in early summer, once the pheromone traps indicate the presence of midge. Phenotyping is carried out by visual inspection of individual florets on at least 10 spikes from 10 independent plants normally four to five weeks after flowering at the milk stage. A plant is noted as resistant if no midge larvae are present within the spikes. Conversely, a plant is susceptible if at least one larva is observed within the sampled spikes.

[0079] The invention also relates to various methods of using the plants of the invention. More particularly the invention relates to a method for conferring or improving resistance to OWBM in a plant, comprising the steps consisting of:

c) transforming at least a plant cell or plant tissue with a vector containing, as a transgene, a nucleic acid construct according to the invention.
d) cultivating the cell(s) or plant tissue thus transformed so as to generate a transgenic plant containing at least a cell which contains, in its genome, at least said nucleic acid construct.

wherein said transgenic plant presents a newly conferred or an improved resistance compared to a plant not comprising said nucleic acid construct.

[0080] A method for identifying (i.e. screening for, selecting) a transgenic plant that can be used in a selection (i.e. breeding) process for obtaining a plant with improved or newly conferred resistance to OWBM, which comprises the step of identifying in a population of plants, the plants containing the nucleic acid construct as described above, is also part of the invention.

[0081] Such method is thus an *in vitro* method, intended to identify, in a population of plants, the ones that carry the transgene according to the invention.

[0082] A breeding process for obtaining a plant with improved or newly conferred resistance is performed by known methods in the art, by crossing, back-crossing and stabilizing plants which present an improved or newly conferred resistance.

[0083] The method for identifying a plant with improved or newly conferred resistance to OWBM, which comprises the step of identifying, in a population of plants, the plants containing the nucleic acid construct or transgene as described above, is also part of the invention. Improved or newly conferred resistance is determined after comparison with isogenic plants which do not contain the nucleic acid construct or transgene.

[0084] A transgenic plant with a newly conferred resistance is intended to mean a plant which is initially susceptible to OWBM and presents a resistant phenotype after acquiring the transgene by transformation or by a breeding process. A transgenic plant with improved resistance is intended to mean a plant which is initially presenting a resistance to OWBM due to the presence of at least one resistant locus to OWBM and presents a more efficient resistance or a more durable resistance after acquiring the transgene by transformation or by breeding process.

[0085] In a specific embodiment, the selection or identification is performed through the use of a marker that is specific to the transgene. In this embodiment, the selection step is thus preferably preceded by a step comprising genotyping said population of plants.

[0086] In a specific embodiment, the selection step is preceded by a step comprising extracting the RNA from the individuals in said population.

[0087] In a specific embodiment, the selection step is preceded by a step comprising extracting proteins from the individuals in said population.

[0088] In a specific embodiment, said population is the progeny obtained from crossing a transgenic plant, wherein said transgene comprises the expression cassette as described above, with a plant line which does not contain said transgene (the recipient plant line).

[0089] In a preferred embodiment, the methods are applied to a cereal, in particular, rice, maize, wheat, barley. It is preferred when said plant is wheat.

[0090] The invention also encompasses markers used to identify the Sm1 resistance locus in plants. The Sm1 locus is intended to mean a chromosomal region genetically linked to and containing the genes for resistance to OWBM. More specifically, the Sm1 locus comprises the chromosomal interval defined by the SNP markers Wa_c6957_32 and Excalibur c1787_1301 (Figure 2), wherein said Wa_c6957_32 marker is SEQ ID NO: 9 or SEQ ID NO: 10, and said Excalibur c1787_1301 marker is SEQ ID NO: 11 or SEQ ID NO: 12, of the IWGSC Whole genome assembly ('IWGSC WGA') Chinese Spring genome. Table 1A and 1B gives details of the DNA sequences used to develop codominant SNP and real-time PCR markers respectively, that were developed by the applicant in the very small genetic interval of 0.067 cM for use in marker-assisted selection for the Sm1 locus.

[0091] As used herein a "marker" refers to a specific DNA sequence identified within the genome of a plant and which can be used to determine whether a plant has inherited a particular phenotype or allele of interest from a parent plant.

[0092] The marker may include coding or non-coding sequences. In particular, said marker may be a fragment of the genomic sequences SEQ ID NO:1 or SEQ ID NO:4 that allow to detect respectively the presence or absence of the nucleic acid encoding for RGA 1 or the presence of the nucleic acid encoding for RGA 2 in a plant. Said marker may also be a fragment of the respective corresponding coding sequences SEQ ID NO:2 or SEQ ID NO:5.

[0093] More specifically the marker used to identify the presence or absence in the genome of a plant of the nucleic acid encoding for RGA 1 is the nucleic acid fragment encoded by the peptide defined by LRR motif of SEQ ID NO:3 said nucleic acid fragment having the sequence as depicted in SEQ ID NO: 7

[0094] More specifically the marker used to identify the presence or absence in the genome of a plant of the nucleic

acid encoding for RGA 2 is the nucleic acid fragment encoded by the peptide defined by LRR motif of SEQ ID NO:6 said nucleic acid fragment having the sequence as depicted in SEQ ID NO: 8.

[0095] In particular the markers of the invention can also be used as a probe to identify and isolate orthologs of genes encoding RGA 1 and RGA 2. Orthologs is intended to mean genes found in different species having a common ancestor and encoding for proteins driving a similar function in the plant that is, in the present invention, resistance to OWBM.

[0096] More particularly, the probe to identify and isolate in a plant genome RGA 1 orthologs is depicted in SEQ ID NO:7 and the probe to identify and isolate RGA 2 orthologs is depicted in SEQ ID NO:8.

[0097] Any method known in the art may be used in the art to assess the presence or absence of a nucleic acid sequence in the genome of a plant. Some suitable methods include, but are not limited to, sequencing, hybridization assays, polymerase chain reaction (PCR), ligase chain reaction (LCR).

[0098] Markers of the invention may also include one or more Single Nucleotide Polymorphism or SNP identified between two different susceptible and resistant genomes. It is also possible to identify sequence deletion/insertion (INDEL) polymorphism.

[0099] Said marker is preferentially localized within the Sm1 locus comprising both RGA 1 and RGA 2 nucleic acids and bounded by the markers Wa_c6957_32 and Excalibur_c1787 (Figure 2), wherein said Wa_c6957_32 marker is SEQ ID NO: 9 or SEQ ID NO: 10, and said Excalibur_c1787_1301 marker is SEQ ID NO: 11 or SEQ ID NO: 12, of the IWGSC Whole genome assembly ('IWGSC WGA') Chinese Spring genome. More preferentially, it is herein disclosed SNPs within markers (identified by their nucleotide sequence) for determining, in a plant, whether any recombinant chromosomal fragment retains the Sm1 allele conferring OWBM resistance. Table 1 provides the DNA sequences of both the susceptible and the resistant alleles in wheat.

Table 1: Marker sequences:

A- Marker showing a single nucleotide polymorphism (SNP) between resistant and susceptible wheat genomes

Marker sequence of resistance allele	Marker sequence of susceptible allele
SEQ ID NO:9 Marker Wa_c6957_32 CGGGACAGCCAAGAGAAATTCCATTTGGCGA[T]CGTTC AAATGTGCACTGCATTCTCGTACTGCCGTCGCCGTCGT CTGTCTCTTGACTGTCG	SEQ ID NO:10 Marker Wa_c6957_32 CGGGACAGCCAAGAGAAATTCCATTTGGCGA[C]CGTT CAAATGTGCACTGCATTCTCGTACTGCCGTCGCCGTCG TCTGTCTCTTGACTGTCG
SEQ ID NO:11 Marker Excalibur_c1787 CATATTGATGAACAAGAACAAGTATAGTGTGCGGCCA AGCCAAGGGACCATGCCACCGTGCTCCAGGCGTTATGT TGTCGTGAC[G]CTGTCAGCGCAAGAGGCGGCGCCGCC ATACATGCGGTGTGACGACATGCTCCTAGTGCAGAGCA CCAGCATCACCCAAGATCTTGGTGAGATCAATTATCAA GAATTGTTCGACGTGGCCAGGGCGGATA	SEQ ID NO:12 Marker Excalibur_c1787_1301 CATATTGATGAACAAGAACAAGTATAGTGTGCGGCCA AGCCAAGGGACCATGCCACCGTGCTCCAGGCGTTATGT TGTCGTGAC[A]CTGTCAGCGCAAGAGGCGGCGCCGCC ATACATGCGGTGTGACGACATGCTCCTAGTGCAGAGCA CCAGCATCACCCAAGATCTTGGTGAGATCAATTATCAA GAATTGTTCGACGTGGCCAGGGCGGATA
SEQ ID NO :13 Marker 1 GTATGAAAAGTATGAAAATAGCACTTGCTTGTATGTAG ACCTACGGTTT[T]CTAACTATAGACTTAGTAAACATACC ACATGAAATAACATACCA	SEQ ID NO :14 Marker 1 TATCAAAAGTATGAAAATAGCACTTGCTTGTATGTAGA CCTACGGTTT[A]CTAACTATAGACTTAGTAATAAACATA CCACATGAAATAACATACCA
SEQ ID NO :15 Marker 2 TGCTCAACAGCTCAAGTACCTTTTATCCTTTAGATGCTC GGTGAGGTCGTGAATGAG[C]TCGTGCACCTCATTAGC GACGGGTGGTTGATCCGGACGAACTTGTGCGAGTATG CTCCTCAGGATCCTCCTCATGTCAGGTTTCTTGGCGGTC CGCACGAAAGCCCGGCAGCAGAAGT	SEQ ID NO :16 Marker 2 TACCTTTTATCCTTCAGATGCTCGGTGAGGTCATGAATG AG[G]TCGTGCACCTCATTAGCGTCGGGTGGTTGGTGC GGACGAACTTGTGCGAGTATGCTCCTCAGGATCCTCCT CATGTCAGGTTTCTTGGCCGTCCGCACAAAAGCTCGGC AGTCGAAGT
SEQ ID NO :17 Marker 3 TGCTCAACAGCTCAAGTACCTTTTATCCTTTAGATGCTC GGTGAGGTCGTGAATGAGCTCGTGCACCTCATTAGCG ACGGGTGGTTG[A]TCCGGACGAACTTGTGCGAGTATG CTCCTCAGGATCCTCCTCATGTCAGGTTTCTTGGCGGTC CGCACGAAAGCCCGGCAGCAGAAGT	SEQ ID NO :18 Marker 3 TACCTTTTATCCTTCAGATGCTCGGTGAGGTCATGAATG AGGTCGTGCACCTCATTAGCGTCGGGTGGTTG[G]TGC GGACGAACTTGTGCGAGTATGCTCCTCAGGATCCTCCT CATGTCAGGTTTCTTGGCCGTCCGCACAAAAGCTCGGC AGTCGAAGT
SEQ ID NO :19 Marker 4 TGCTCAACAGCTCAAGTACCTTTTATCCTTTAGATGCTC GGTGAGGTCGTGAATGAGCTCGTGCACCTCATTAGCG ACGGGTGGTTGAT[C]CGGACGAACTTGTGCGAGTATG CTCCTCAGGATCCTCCTCATGTCAGGTTTCTTGGCGGTC	SEQ ID NO :20 Marker 4 TACCTTTTATCCTTCAGATGCTCGGTGAGGTCATGAATG AGGTCGTGCACCTCATTAGCGTCGGGTGGTTGGT[G]C GGACGAACTTGTGCGAGTATGCTCCTCAGGATCCTCCT CATGTCAGGTTTCTTGGCCGTCCGCACAAAAGCTCGGC

Marker sequence of resistance allele	Marker sequence of susceptible allele
CGCACGAAAGCCCGGCAGCAGAAGT	AGTCGAAGT
SEQ ID NO :21 Marker 5 CCAAGTGTTAGTATACTCTAGGAAGCTTTATGCGCCAA CTTTGCATGTAGGTAACTAAAAACTGGG[T]GAAAAGG TTGCTGACAGAGCCTGGATATTCCTCGCACGAAACCGA CATGCTAGGAGTCACTAGATGTGTTGAAATTG	SEQ ID NO :22 Marker 5 CCAAGTGTTAGTATACTCTAGGAAGCTTTATGCGCCAA CTTTGCATGTAGGTAACTAAAAACTGGG[C]GAAAAGG TTGCTGACAGAGCCTGGATATTCCTCGCACGAAACCGA CATGCTAGGAGTCACTAGATGTGTTGAAATTG

B- Marker showing a presence or absence in the plant genome (Real-Time PCR marker in the RGA genes)

	Primer F	Primer R	Probe MGB*	Amplicon or Marker sequence
NBS1 8000	SEQ ID NO :23 CGACGGCGCACG TGAT	SEQ ID NO :24 AACGGACGACGAAT GCAAAT	SEQ ID NO :25 TGGACTCGATCCATTG	SEQ ID NO :26 CGACGGCGCACGTGATGCTGGAC TCGATCCATTGCATTTGCATTCGTC GTCCGTT
NBS4 4000	SEQ ID NO :27 ACGCTCCGCAAAA ATCTGA	SEQ ID NO :28 AACAACGAATAGCGC CTTGAG	SEQ ID NO :29 TGACCGGGACGAGCA	SEQ ID NO :30 ACGCTCCGCAAAAATCTGATTGAC CGGGACGAGCAGCTCAAGGCGCT ATTCGTTGTT

**MGB : Minor Groove Binder*

[0100] As used herein, the term "allele(s)" means any of one or more alternative forms of a gene at a particular locus. In a diploid (or amphidiploid cell of an organism), alleles of a given gene are located at a specific location or locus on a chromosome. One allele is present on each chromosome of the pair of homologous chromosomes.

[0101] Whenever reference to a "plant" or "plants" is made, it is understood that also plant parts (cells, tissues or organs, seed pods, seeds, severed parts such as roots, leaves, flowers, pollen, etc.), progeny of the plants which retain the distinguishing characteristics of the parents. "Plant" also relates to a line, a variety, a cultivar or an elite plant.

[0102] The invention relates to a method for identifying a plant resistant to OWBM comprising the Sm1 locus. The method encompasses the identification of a OWBM resistant plant compared to a plant lacking the nucleic acid of the present invention wherein the method is comprising the steps of:

a) isolating DNA genomic sequence or RNA nucleic acid from the plant, and
b) identifying the presence of the nucleic acid of the invention or the presence of resistant alleles and/or the absence of susceptible alleles of a marker within the Sm1 locus as defined above.

[0103] In a preferred embodiment, the methods are applied to a cereal, in particular, rice, maize, wheat, barley. It is preferred when said plant is wheat.

[0104] The presence or absence of the nucleic acid of the invention is identified with a marker or a fragment derived from said nucleic acid. The method can be based on the identification of the presence of either one nucleic acid encoding either RGA 1 or RGA 2, or the presence of both nucleic acid encoding RGA 1 and RGA 2 in a resistant plant. Similarly, the method will be based on the identification of the absence of the such nucleic acids in the susceptible plant.

[0105] Markers of interest within the Sm1 locus according to the present invention, defined by the SNPs Wa_c6957_32 and Excalibur_c1787 (Figure 2), wherein said Wa_c6957_32 marker is SEQ ID NO: 9 or SEQ ID NO: 10, and said Excalibur c1787_1301 marker is SEQ ID NO: 11 or SEQ ID NO: 12, of the IWGSC Whole genome assembly ('IWGSC WGA') Chinese Spring genome, are linked to Sm1 resistance. The expression "markers linked to Sm1 resistance" means markers showing a DNA polymorphism between a resistant plant and a susceptible plant or markers amplifying a DNA region from the resistant plant, which is absent in the susceptible plant. Preferentially the markers within the Sm1 locus are the markers as listed in Table 1. More preferentially, the markers of interest are the codominant markers listed in Table 1A. Codominant markers are more informative, as they can detect heterozygous plants which have inherited both a resistant and susceptible allele and more accurate, as a failed PCR amplification of the marker would not be interpreted as being susceptible as could be the case for a dominant marker targeting the resistance allele.

[0106] Any method known in the art may be used in the art to assess the presence or absence of a SNP. Some suitable methods include, but are not limited to, sequencing, hybridization assays, polymerase chain reaction (PCR), ligase chain reaction (LCR), and genotyping-by-sequence (GBS), or combinations thereof.

[0107] Different PCR-based methods are available to the person skilled of the art. One can use the Real-Time PCR method or the KASP method from KBioscience (LGC Group, Teddington, Middlesex, UK).

[0108] Successive rounds of PCR amplifcation are performed to amplify the target sequence (marker) using fluorescently labelled primers. The nature of the emitted fluorescence is used to identify the allelic form or forms (as in the case of a heterozygote) present in the mix from the studied DNA.

[0109] As used herein, a primer encompasses any nucleic acid that is capable of priming the synthesis of a nascent nucleic acid in a template-dependent process, such as PCR. Typically, primers are oligonucleotides from 10 to 30 nucleotides, but longer sequences can be employed. Primers may be provided in double-stranded form though single-stranded form is preferred. Alternatively, nucleic acid probe can be used. Nucleic acid probe encompasses any nucleic acid of at least 30 nucleotides and which can specifically hybridize under standard stringent conditions with a defined nucleic acid. Standard stringent conditions as used herein refers to conditions for hybridization described for example in Sambrook et al. (1989) which can comprise 1) immobilizing plant genomic DNA fragments or library DNA on a filter 2) pre-hybridizing the filter for 1 to 2 hours at 65°C in 6x SSC 5x Denhardt's reagent, 0.5% SDS and 20mg/ml denatured carrier DNA 3) adding the probe (labeled) 4) incubating for 16 to 24 hours 5) washing the filter once for 30min at 68°C in 6x SSC, 0.1% SDS 6) washing the filter three times (two times for 30min in 30ml and once for 10 min in 500ml) at 68°C in 2x SSC 0.1% SDS.

[0110] It is also presently disclosed but not as part of the invention, a method for introgressing the *Sm1* resistance into a plant, the method comprising the steps of:

a) crossing a first resistant plant comprising the nucleic acid of the invention or a resistant allele as listed in Table 1 with a second plant wherein said second plant is not comprising the nucleic acid of the invention or a resistant allele as listed in Table 1
b) backcrossing the progeny of step a) with said second plant
c) identifying and selecting from b) one or more plants comprising the nucleic acid of the invention or a resistant allele as listed in Table 1 associated with OWBM resistance.

[0111] As used herein, the term "introgression" refers to a DNA fragment of a particular plant transferred into a second plant. More specifically, in a first purpose said second plant is a susceptible plant to OWBM. Introgression allows the production of a new resistant plant from a susceptible one by transferring from a resistant plant a chromosomal fragment comprising the nucleic acid of the invention to a susceptible plant. The transfer of said nucleic acid is identified with the markers of the present invention.

[0112] In a second purpose, said second plant is resistant to OWBM due to the presence in its genome of at least one resistance locus different from Sm1 locus. The introgression of Sm1 locus in said second plant allows the combination of Sm1 locus with other resistance loci to OWBM. This process is well-known as "gene pyramiding" or 'gene stacking". In the case of resistance genes, the stacking improves the resistance by providing a plant with increased durability of resistance when compared to a plant having only one resistance locus.

[0113] In a preferred embodiment, the methods are applied to a cereal, in particular, rice, maize, wheat, barley. It is preferred when said plant is wheat.

[0114] The invention also relates to the use of the nucleic acid of the invention or any fragment of said nucleic acid as a screening tool for identifying a new ligand peptide interacting with the protein encoded by nucleic acid of the invention.

[0115] Methods of the art are well known for the identification of ligand-protein interaction like, for example, the yeast two-hybrid system (Fields and Song, 1989) or by immunoprecipitation.

[0116] It is also presently disclosed but not as part of the invention, a method for modifying the nucleic acid of the invention to improve the function of the protein encoded by said nucleic acid in order to improve the resistance to OWBM.

[0117] Method of the art related to the modification of the genomic DNA or "gene editing" are well-known like, for example TALENs (WO2011072246) or CRISPR Cas9 (WO2013181440).

[0118] The present application also provides the promoter regions of the genes encoding the proteins conferring resistance to OWBM according to the present invention and encompasses a sequence with at least 80%, preferably with at least 85%, more preferably with at least 90%, and even more preferably with at least 95% sequence identity with the sequence of these promoter regions. The identity percentage is calculated by any sequence analysis method known by the skilled person, and particularly with algorithms such as Needleman-Wunsch. The identity percentage is calculated over the whole length of the query sequence.

[0119] In a preferred embodiment, the sequences SEQ ID NO: 31 or SEQ ID NO: 32 or any sequence with at least 80%, preferably with at least 85%, more preferably with at least 90%, and even more preferably with at least 95% sequence identity with SEQ ID NO: 31 or SEQ ID NO: 32 comprises the promoter region of the gene encoding a protein

conferring resistance to OWBM according to the invention. In an advantageous embodiment of the present invention, the sequences SEQ ID NO: 33 to SEQ ID NO: 37 or any sequence with at least 80%, preferably with at least 85%, more preferably with at least 90%, and even more preferably with at least 95% sequence identity with SEQ ID NO: 33 to SEQ ID NO: 37 comprises the promoter region of the gene encoding a protein conferring resistance to OWBM according to the invention.

EXAMPLES:

Example 1. Fine-mapping of the genomic region containing the Sm1 locus

A: Fine-Mapping based on genetic linkage maps:

[0120] The Sm1 gene for resistance to Orange Wheat Blossom Midge (OWBM) was first mapped to the short arm of chromosome 2B by Thomas et al (2005) using SSR markers and its map position has recently been refined by Kassa et al (2016). However, the populations used for the linkage mapping were relatively small and so the resolution of these genetic maps is not very high. As a consequence the inventors had to develop a large F2 population and additional DNA markers to saturate the region in order to screen for new recombination events. A total of 259 new SNPs have been developed for the analysis of recombinant plants, as described below, in order to fine map the region of interest and to identify the genes at the Sm1 resistance locus.

[0121] In 2011, 4,466 F2 plants were generated from selfing an F1 cross made between the UK winter wheat varieties Xi19 (susceptible) and Robigus (Sm1 resistant). These individuals were genotyped with eight SNP markers that were evenly distributed between 12.3 cM and 41.3 cM on the genetic map of chromosome 2B to identify recombinants across this region. A subset of 376 lines were further genotyped with an additional set of 42 SNP markers in order to produce a more detailed genetic map of the region around the Sm1 locus.

[0122] Following two additional rounds of selfing of the recombinant F2 plants with marker selection, a total of 111 F4 homozygous recombinants were generated across the Sm1 region on 2BS.These F4 recombinants were genotyped with the additional 42 SNP markers described above and phenotyped in during the summer of 2013 for resistance to OWBM under field conditions. Phenotyping was carried out, four to five weeks after flowering at the milk stage (Kassa et el, 2016; Ellis et al, 2009), by visual inspection of individual florets on at least 10 spikes from 10 independent plants per F4 family. F4 families with a complete absence of midge larvae within the spikes were scored as being resistant; whereas if a single larvae was found then the whole F4 family was scored as being susceptible. Midge infestation levels were good and the Sm1 locus was mapped, without any ambiguity, to a genetic interval of 0.49 cM (defined by 22 recombinants out of the 4466 F2 plants).

[0123] By mining the available in-house and public genomic resources, an extra 186 SNPs were identified within or closely flanking this 0.49 cM interval. These extra markers allowed further discrimination of the 22 recombinant plants and reduced the genetic interval containing Sm1 to just 0.067 cM. Further analysis of the physical map (see Example 1: B) and sequences from the transcriptomics experiment (see Example 3) added 23 extra SNPs to the region which gave a total of 259 markers tightly-linked to the Sm1 locus.

B: Fine-mapping based on physical maps:

[0124] Both the IWGSC chromosome survey contigs (The International Wheat Genome Sequencing Consortium (IWGSC) (2014)) and the IWGSC Whole genome assembly, 'IWGSC WGA' (available from June 2016 from the URGI IWGSC repository) were used as the genomic resources to enable the physical mapping of the Sm1 locus.

[0125] The right border of the Sm1 interval was anchored to the IWGS contig 2BS:5245756 (11643 bp) with 21 SNP markers on this contig. Three SNP markers (e.g. Excalibur c1787_1301 depicted in SEQ ID NO:10) confirmed the recombination breakpoint on the IWGS contig 2BS:5245756. The left border was anchored to the IWGS contig 2BS:5191992 (937 bp) by a single SNP (i.e. Wa_c6957_32 depicted in SEQ ID NO:9).

[0126] Following the release of the 'IWGSC WGA' wheat genomic sequence, a total of 93 SNP markers were anchored, to the region of interest, on two scaffolds: IWGSCWGAV02_2BS_scaffold14096 and IWGSCWGAV02_2BS_scaffold11627 that are 675,277 bp and 2,720,717 bp in length respectively. The genetic order of 93 SNP markers in the Xi19 x Robigus bi-parental mapping population was consistent with the physical position of the SNPs in the 'IWGSC WGA' scaffolds. The SNP markers Wa_c6957_32 and Excalibur_c1787_1301 delimit the borders of the Sm1 interval (Figure 2) and define a region of 941,325 bp on the 'IWGSC WGA' Chinese Spring genome (122,033 bp on IWGSCWGAV02_2BS_scaffold14096 and 819,292 bp on IWGSCWGAV02_2BS_scaffold11627; Figure 2).

[0127] In parallel, a Renan BAC library, available at the INRA GNRG Plant Genomic Center (http://cnrgv.toulouse.inra.fr/), was also screened. Renan is a midge resistant, French winter wheat variety, which is identical to Robigus at the

Sm1 locus, as determined by genotyping both varieties with the 21 SNPs developed within the interval defined by Wa_c6957_32 and Excalibur_c1787_1301. The BAC library was screened with four SNP markers located on the 2BS:5245756 contig, plus one SNP from each of the five IWGSC contigs in the vicinity of 2BS:5245756 (i.e. 2BS:5157057, 2BS:5163033, 2BS:5191992, 2BS:5174837 and 2BS:5175242).

[0128] One BAC clone (715D09) of 110 kbp in length was identified which contained the right border of the Sm1 locus. Annotation of this BAC clone revealed two disease resistance gene analogues: RGA 1 and RGA 2 and a cysteine-rich receptor kinase gene. Real-Time PCR markers developed from the sequence of the two RGA genes (Table 1B) were genotyped on a subset of recombinant plants from the Xi19 x Robigus mapping population, which confirmed their genetic position within the interval.

[0129] RGA 1 gene is encoding a protein of sequence SEQ ID NO: 3 and RGA 2 gene is encoding a protein of sequence SEQ ID NO: 6. Both proteins have the CC, NBS, and LRR motifs of a classical CNL resistance genes as shown on Figure 1. Interestingly the annotation of RGA 2 revealed the presence of extra domains that can also be involved in plant defense (i.e. a NAM, pKinase and a PapD-like domains). Recent studies argue that these "integrated domains" may arise from fusions between NLRs (Nucleotide-binding Leucine-rich Repeats) and that these additional domains serve as "baits" for the pathogen-derived effector proteins; thus enabling pathogen recognition (Sarris et al, 2016).

[0130] BLAST analysis of the two Robigus RGA cDNA's against the whole IWGSC WGA dataset identified a region within IWGSCWGAV02_2BS_scaffold11627 with a percentage of identity ranging from 78 to 82 % for fragments larger than 700 bp. More specifically, the 1,173 and 968 nucleotides coding for the LRR domains of RGA 1 and RGA 2 respectively, have two BLAST hits each both in an interval of 141,193 bp within IWGSCWGAV02_2BS_scaffold11627 (i.e. RGA.CS1 and RGA.CS2 shown in Figure 2). The percentages of identity of 77, 81 and 77, 27 for RGA 1 and of 81, 75 and 80, 75 for RGA 2 suggest a tandem duplication of the two ancestral RGA genes. The observed percentage of identity did not differ significantly from those obtained later by BLASTing mRNA contigs against the whole IWGSC WGA (see section entitled "Differential expression on recombinant plants"). No higher percentages of identity were identified elsewhere in the genome suggesting a co-ancestry between RGA 1 and RGA 2 and the cluster of RGA genes identified within IWGSCWGAV02_2BS_scaffold11627 (i.e. RGA.CS1 and RGA.CS2 shown in Figure 2). The IWGSCWGAV02_2BS_scaffold11627 fragment, remaining within the Sm1 interval, was annotated based on the MAKER score and the ST/mRNAseq data was used to check the quality of the annotation. Only high confidence (HC) genes supported by functional annotation (Uniprot/Swissprot) were retained.

[0131] The 390 bp of the pkinase domain of RGA 2 matched with a percentage of identity of 96.69% to a kinase located at the recombination breakpoint on the IWGSCWGAV02_2BS_scaffold11627. The co-dominant markers listed in Table 1A are located in this pKinase domain. The cysteine-rich receptor kinase at the end of the BAC has been mapped outside the interval (Figure 2).

[0132] Percentages of identity to other kinases present in the IWGSCWGAV02_2BS_scaffold11627 were lower than 90%. These data suggest an ancestral re-arrangement of this region in the midge resistant lines that removed nearly 600 kb between the first two Chinese Spring RGA genes (i.e. RGA.CS1 and RGA.CS2 shown in Figure 2) and this kinase.

[0133] It should be noted that Kassa et al (2016) describe Bradi5g00870 as a putative candidate gene for Sm1, but this RGA gene sequence is not found within the target interval identified in the present invention.

Example 2: Sequencing of Robigus and Xi19

[0134] A 30x genome coverage of both Xi19 and Robigus was generated. Bio-informatic mapping of the gDNA reads to the IWGSCWGAV02_2BS_scaffold14096, IWGSCWGAV02_2BS_scaffold11627 and the BAC sequence confirmed large structural differences in the region, which explain the observed lack of recombination. Essentially, there were no gDNA-seq reads from Xi19 that mapped onto the BAC sequence of Renan (Figure 3). Similarly, none of the gDNA-seq reads from Robigus mapped onto scaffold11627 portion located in the interval. This indicates that the gene content of the corresponding genomic regions in susceptible and resistant plants is very divergent and this was probably caused by a local chromosomal rearrangement. Similar structural rearrangements between the Chinese Spring reference genome sequence and the region containing the Fusarium head blight resistance gene (Fhb1) have recently been published (Schweiger et al, 2016).

Example 3: Differential expression on recombinant plants

[0135] Due to the major structural differences between 'Chinese Spring' reference genome and Renan at the Sm1 locus, the next approach taken was to sequence the transcriptome from wheat spikes taken from 12 resistant and 13 susceptible plants that had undergone a recombination event very close to the Sm1 locus, as well as the two parental lines: Xi19 and Robigus. The objective was to gather expression evidence for the candidate genes identified within the region and to check for the presence of new genes that were absent from the reference genome, but present at the Sm1 locus. The 25 selected plants, plus the parents, were sown in the field and individual spikes were sampled on dry ice,

following midge attack. The mRNA was extracted from these 25 tissue samples and used for paired-end sequencing using an Illumina HiSeq 2500 sequencer, which was performed by external service provider: Service XS. The RNA-Seq Illumina data sets generated on the 25 lines were analysed to produce a de novo transcriptome assembly. The RNA-Seq reads from each individual line were then mapped back onto this de novo transcriptome assembly and the gene expression of the 12 resistant lines was compared to that of the 13 susceptible lines in order to identify differentially expressed genes.

[0136] A first screen for differentially expressed sequences, highlighted the presence of 566 genes that were annotated as insect genes and these were only found in the susceptible plant samples; thus confirming that the sampled wheat spikes had contained developing midge larvae. Therefore the absence of insect gene expression in the resistant plant samples can be interpreted as the direct effect of the Sm1 resistance locus.

[0137] Amongst the differentially expressed plant genes, six were over-expressed sequences in the resistant lines and 21 were over-expressed in the susceptible lines. The sequence of the 27 assembled EST contigs was used for BLAST analysis against both the IWGSC WGA and the BAC sequence to validate their genomic position (Table 2). Only six genes (3 over-expressed and 3 under-expressed genes), out of the total of twenty-seven, mapped to within the target interval (Table 2). The rest of contigs had BLAST hits with percentages of identity >97% to regions on other wheat chromosomes or were physically distant on the chromosome 2B (Table 2).

[0138] Table 2 below shows the chromosomal location of the 27 differentially expressed ESTs identified between the resistant and susceptible plants subjected to RNA-Seq analysis. Below "+" *indicates over-expression in resistant lines and "-" indicates over-expression in susceptible lines.* The IWGSC WGA hits on 2BS for EST 1, 3 and 6 identified IWGSC WGA contig IWGSCWGAV02_2BS_scaffold11627. EST21 (2BS*) maps more than 200 Mbps from the Sm1 locus. EST22 (RGA.Xi19) has been annotated as an RGA protein and it has 89% percentage of identity to RGA.CS1 (Figure 2).

Feature	Expression		IWGSC WGA	Identity percentage with IWGSC WGA	Identity percentage with BAC 715D09
EST1	+	RGA 1	2BS	78%	100%
EST2	+		5DL	100%	
EST3	+	RGA 2	2BS	87%	100%
EST4	+		5BS	98%	
EST5	+		5DL	100%	
EST6	+	Kinase	2BS	77%	100%
EST7	-		6AL	100%	
EST8	-		7DL	96%	
EST9	-		2AL	99%	
EST10	-		1DS	98%	
EST11	-		4DL	99%	
EST12	-		1AL	97%	
EST13	-		3AL	96%	
EST14	-		7DL	99%	
EST15	-		6BS	89%	
EST16	-		2DL	97%	
EST17	-		6DL	98%	
EST18	-		2BL	99%	
EST19	-		3DL	100%	
EST20	-		4DL	100%	
EST21	-		2BS*	99%	
EST22	-	RGA.Xi19	2BS	89%	
EST23	-		6DL	100%	
EST24	-		3DL	98%	
EST25	-		4AS	98%	
EST26	-		2BS	100%	
EST27	-		2BS	100%	

[0139] As Sm1 resistance has been reported as being partially dominant (McKenzie et al, 2002), the following analyses

focused on the six over-expressed genes. Amongst these candidates (EST1 to 6 in Table 2), three genes were located on the BAC clone (715D09). As expected, these three genes were absent from Xi19 and so it was not possible to map any of the Xi19 RNA-Seq reads to the BAC sequence. (N.B. EST1 corresponds to RGA 1 and EST 3 corresponds to RGA 2). And conversely, it was not possible to map any of the Robigus RNA-Seq reads to the IWGSCWGAV02_2BS_scaffold11627 fragment that remained in the interval. The other three genes mapped, with strong BLAST hits, to chromosomes 5DL and 5BS.

Example 4: Marker analysis around the Sm1 locus using a diverse panel

[0140] In total, 23,400 lines were screened from crosses segregating for Sm1 with ten SNP markers distributed across the Sm1 locus in order to search for extra recombinants within this region. In total, 576 putative recombinants were identified and these, plus their parental lines, were also genotyped with two Real-Time PCR markers developed from RGA 1 and RGA 2 (Table 1B) and 46 markers that included 24 SNPs from within the interval and 22 markers tightly flanking the region. The results showed that no recombination events were found between the two RGA genes and no recombinant plants were found within the small 0.067 cM region identified in the Xi19 x Robigus bi-parental mapping population. The lack of recombination within the region is due to the absence of any sequence homology between resistant and susceptible lines. Moreover, all the lines that carried the two Robigus RGA genes shared the Robigus haplotype based on the 24 markers within the target interval suggesting a single origin and a common ancestor for the Sm1 resistance locus.

[0141] Amongst the 576 lines, a sub-panel of 113 diverse lines was selected for phenotypic analysis. This sub-panel contained many recombinant plants arising from different genetic origins in order to validate any potential diagnostic SNP markers. All the recombinant plants and their parental lines were sown and genotyped in summer 2014. The presence of the two RGA genes was always shown to be 100% diagnostic for the presence of Sm1. From these 24 markers within the interval, five were found to be correlated with the presence and absence of the two RGAs (Table 1A), which makes them ideal for marker-assisted selection of the Sm1 gene.

Example 5: Identification of genomic regions that have promoter motifs for RGA 1 and RGA 2 genes

[0142] The 715D09 BAC sequence containing the two RGA genes was submitted to the TSSP (Prediction of PLANT Promoters (Using RegSite Plant DB, Softberry Inc.)) using the algorithm as implemented on the www.softberry.com website.

[0143] The sequences of the arbitrary 3000 bp upstream of the translation start site (TSS) of the two genes of interest containing the predicted promoter sequences are as set forth in SEQ ID NO: 31 for RGA 1 and SEQ ID NO: 32 for RGA 2. Within those 3000 bp, for both genes, extra predicted promoter regions containing TATA box and transcription factor binding sites were identified.

[0144] The positions of predicted promoter sequences in SEQ ID NO: 31 based on TSSP algorithm are the following:

- Promoter Position 285 with TATA box at position 251;
- Promoter Position 1380 with TATA box at position 1364;
- Promoter Position 2609 with TATA box at position 2591.

[0145] The positions of predicted promoter sequences in SEQ ID NO: 32 based on TSSP algorithm as the following:

- Promoter Position 942 with TATA box at position 907;
- Promoter Position 2475 with TATA box at position 2461.

Example 6: Functional study of RGA 1 and RGA 2 genes

[0146] Both RGA 1 and RGA 2 gene functions can be validated with different methods well known in the art. Genetic transformation of a susceptible wheat cultivar overexpressing RGA 1 or RGA 2 under different promoters can be obtained and tested for their ability to confer OWBM resistance in glass-house conditions or in the field. Validation can also be achieved by mutagenesis with methods known from skilled person in the art, with for example, EMS treatment. The validation consists of obtaining several independent "loss-of-resistance" mutants derived from the EMS treatment of a resistant wheat cultivar and further identifying mutations within the candidate gene; thus confirming the resistance function of the gene. For example, such method is described by Periyannan et al. (2013) used to identify the wheat stem rust resistance gene Sr33.

REFERENCES:

[0147]

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402

Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, Yu YK (2005) Protein database searches using compositionally adjusted substitution matrices. The FEBS Journal 272: 5101-5109

Anderson OD, Greene FC (1989) The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of hexaploid bread wheat. Theoretical and Applied Genetics 77: 689-700

Anon (2016) Orange wheat blossom midge. AHDB Information sheet N° 53 Summer 2016 (https://cereals.ahdb.org.uk/)

Bailey TL and Elkan C (1994) Fitting a mixture model by expression maximization to discover in biopolymers. Proceedings of the International Conference on Intelligent Systems for Molecular Biology_ 2: 28-36

Birkett MA, Bruce TJA, Martin JL, Smart LE, Oakley J, Wadhams LJ (2004) Responses of female orange wheat blossom midge to wheat panicle volatiles. Journal of Chemical Ecology 30: 1319-1328

Blake NK, Stougaard RN, Weaver DK, Sherman JD, Lanning SP, Narouka Y, Xue Q, Martin JM, Talbert LE (2011) Identification of a quantitative trait locus for resistance to Sitodiplosis mosellana (Géhin), the orange wheat blossom midge, in spring wheat. Plant Breeding 130: 25-30

Bruce TJ, Hooper AM, Ireland L, Jones OT, Martin JL, Smart LE, Oakley J, Wadhams LJ (2007) Development of a pheromone trap monitoring system for orange wheat blossom midge, Sitodiplosis mosellana, in the UK. Pest Management and Science 63: 49-56

Chavalle S, Censier F, San Martin y Gomez G, De Profta M (2015) Protection of winter wheat against orange wheat blossom midge, Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae): efficacy of insecticides and cultivar resistance. Pest Management Science 71: 783-790

Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Research 5: 213-218

Depigny-This D, Raynal M, Aspart L, Delseny M, Grellet F (1992) The cruciferin gene family in radish. Plant Molecular Biology 20: 467-479

Ding H, Lamb RJ (1999) Oviposition and larval establishment of Sitodiplosis mosellana (Diptera: Cecidomyiidae) on wheat at different growth stages. The Canadian Entomologist 131: 475-481

Ding H, Lamb RJ, Ames N (2000) Inducible production of phenolic acids in wheat and antibiotic resistance to Sitodiplosis mosellana. Journal of Chemical Ecology 26: 969-985

Doane JF, Olfert OO, Elliott RH, Hartley S, Meers S (2013) Sitodiplosis mosellana (Géhin), orange wheat blossom midge (Diptera: Cecidomyiidae). In "Biological control programmes in Canada 2001-2012". Eds PG Mason & DR Gillespie. Chapter 39, pp 272-276

Dogimont C, Chovelon V, Pauquet J, Boualem A, Bendahmane A (2014) The Vat locus encodes for a CC-NBS-LRR protein that confers resistance to Aphis gossypii infestation and A. gossypii-mediated virus resistance. The Plant Journal 80: 993-1004

Du B, Zhang W, Liu B, Hu J, Wei J, Shi Z, He R, Zhu L, Chen R, Han B, He G (2009) Identification and characterization of Bph14, a gene conferring resistance to brown planthopper in rice. Proceedings of the National Academy of Sciences USA 106: 22163-22168

Elliot RH, Mann LW (1996) Susceptibility of red spring wheat, Triticum aestivum L. cv. Katepwa, during heading and anthesis to damage by wheat midge, Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae). The Canadian Entomologist 128: 367-375

Ellis SA, Bruce TJA, Smart LE, Martin JA, Snape J, Self M (2009) Integrated management strategies for varieties tolerant and susceptible to wheat blossom midge. HGCA project report No. 451.

Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. Nature 340: 245-246

Gharalari AH, Fox SL, Smith MAH, Lamb RJ (2009a) Ovipoistion deterrence in spring wheat, Triticum aestivum, against orange wheat blossom midge, Sitodiplosis mosellana: implications for inheritance of deterrence. Entomologia Experimentalis et Applicata 133: 74-83

Gharalari AH, Fox SL, Smith MAH, Lamb RJ (2009b) The relationship between morphological traits of the spring wheat spike and oviposition deterrence to orange wheat blossom midge. Entomologia Experimentalis et Applicata 132: 182-190

Gharalari AH, Smith MAH, Fox SL,Lamb RJ (2011) Volatile compounds from non-preferred wheat spikes reduce oviposition by Sitodiplosis mosellana. The Canadian Entomologist 143: 388-391

Gries R, Gries G, Khaskin G, King S, Olfert OO, Kaminski LA, Lamb R, Bennett R (2000) Sex pheromone of orange wheat blossom midge, Sitodiplosis mosellana. Naturwissenschaften 87: 450-454.

Harris MO, Stuart JJ, Mohan M, Nair S, Lamb RJ, Rohfritsch O (2003) Grasses and gall midges: Plant defense and insect adaptation. Annual Review of Entomology 48: 549-577

Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T (1996) High efficiency transformation of maize (Zea mays L.) mediated by Agrobacterium tumefaciens. Nature Biotechnology, 14: 745-750

Jouanin L, Guerche P, Pamboukdjian N, Tourneur C, Casse Delbart F, Tourneur J (1987) Structure of T-DNA in plants regenerated from roots transformed with Agrobacterium rhizogene strain A4. Molecular and General Genetics 206: 387-392

Kassa MT, Haas S, Schliephake E, Lewis C, You FM, Pozniak CJ, Kramer I, Perovic D, Sharpe AG, Fobert PR, Koch M, Wise IL, Fenwick P, Berry S, Simmonds J, Hourcade D, Senellart P, Duchalais L, Robert O, Forster J, Thomas JB, Friedt W, Ordon F, Uauy C, McCartney CA (2016) A saturated SNP linkage map for the orange blossom midge resistance gene Sm1. Theoretical and Applied Genetics 129: 1507-1517

Kay R, Chan, A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236:1299-1302

Lamb RJ, McKenzie RIH, Wise IL, Barker PS, Smith MAH (2000) Resistance to Sitodiplosis mosellana (Diptera: Cecidomyiidae) in spring wheat (Gramineae). The Canadian Entomologist 132: 591-605

Lamb RJ, Smith MAH, Wise IL, Clarke P, Clarke J (2001) Oviposition deterrence to Sitodiplosis mosellana (Diptera: Cecidomyiidae): a source of resistance for durum wheat (Gramineae). The Canadian Entomologist 133: 579-591

Lamb RJ, Wise IL, Smith MAH, McKenzie RIH, Thomas J, Olfert OO (2002) Oviposition deterrence against Sitodiplosis mosellana (Diptera: Cecidomyiidae) in spring wheat (Gramineae). The Canadian Entomologist 134: 85-96

Lamb RJ, Sridhar P, Smith MAH, Wise IL (2003) Oviposition preference and offspring performance of a wheat midge Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae) on defended and less defended wheat plants. Environmental Entomology 32: 414-420

Lamb RJ, Smith MAH, Wise IL, McKenzie RIH (2015) Resistance to wheat midge (Diptera: Cecidomyiidae) in winter wheat and the origins of resistance in spring wheat (Poaceae). The Canadian Entomologist 1: 1-10.

McElroy D, Zhang W, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. The

Plant Cell 2: 163-171

McKenzie RIH, Lamb RJ, Aung T, Wise IL, Barker P, Olfert OO (2002) Inheritance of resistance to wheat midge, Sitodiplosis mosellana, in spring wheat. Plant Breeding 121: 383-388

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. The Plant Journal 20: 317-332.

Miller BS, Halton P (1960) The damage to wheat kernels caused by the wheat blossom midge (Sitodiplosis mosellana). Journal of Science, Food & Agriculture 12: 391-398

Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods in Enzymology 155:335-350

Oakley JN (1994) Orange wheat blossom midge: a literature review and survey of the 1993 outbreak. Research Review No. 28, HGCA, Hamlyn House, Highgate Hill, London, UK

Oakley JN, Talbot G, Dyer C, Self MM, Freer JBS, Angus WJ, Barrett JM, Feuerhelm G, Snape J, Sayers L, Bruce TJA, Smart LE, Wadhams LJ (2005) Integrated control of wheat blossom midge: variety choice, use of pheromone traps and treatment thresholds. HGCA Project, Report 363

Olfert O, Elliott RH, Hartley S (2009) Non-native insects in agriculture: strategies to manage the economic and environmental impact of wheat midge, Sitodiplosis mosellana, in Saskatchewan. Biological Invasions 11: 127-133

Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, Deal K, Luo M, Kong X, Bariana H, Mago R, McIntosh R, Dodds P, Dvorak J, Lagudah E (2013). The gene Sr33, an ortholog of barley Mla genes, encodes the wheat stem rust race Ug99. Science, 10.1126/ 1239028.

Robert LS, Thompson RD, Flavell RB (1989) Tissue-specific expression of a wheat high molecular weight glutenin gene in transgenic tobacco. The Plant Cell 1: 569-578

Rossi M, Goggin FL, Milligan SB, Klaoshian I, Ullman DE, Williamson VM (1998) The nematode resistance gene Mi of tomato confers resistance against the potato aphid. Proceedings of the National Academy of Sciences USA 95: 9750-9754

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sarris PF, Cevik V, Dagdas G, Jones JDG, Krasileva KV (2016) Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. BMC Biology 14:8

Schweiger W, Steiner B, Vautrin S, Nussbaumer T, Slegwart G, Zamini M, Jungreithmeier F, Gratl V, Lemmens M, Mayer KFX, Berges H, Adam G, Buerstmayr H (2016) Suppressed recombination and unique candidate genes in the divergent haplotype encoding Fhb1, a major Fusarium head blight resistance locus in wheat. Theoretical and Applied Genetics 129: 1607-1623

Sekhwal KM, Pingchuan L, Lam I, Wang X, Cloutier S, You FM (2015) Disease resistance gene analogs (RGAs) in Plants. The International Journal of Molecular Sciences 16: 19248-19290

Smith CM, Clement SL (2012) Molecular bases of plant resistance to arthropods. Annual Review of Entomology 57: 309-28

Stemmer PC (1994) DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proceedings of the National Academy of Sciences USA 91: 10747-10751

Till BJ, Colbert T, Tompa R, Enns LC, Codomo CA, Johnson JE, Reynolds SH, Henikoff JG, Greene EA, Stein MN, Comai L, Henikoff S (2003) High-throughput TILLING for functional genomics. Plant Functional Genomics: Methods

and Protocols. Edited by: Grotewald E. Clifton, NJ, Humana Press, 205-220

The International Wheat Genome Sequencing Consortium (IWGSC) (2014) A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science 345, Issue 6194

Thomas J, Fineberg N, Penner G, McCartney C, Aung T, Wise I, McCallum B (2005) Chromosome location and markers of Sm1: a gene of wheat that conditions antibiotic resistance to orange wheat blossom midge. Molecular Breeding 15: 183-192

Verdaquer B, de Kochko A, Beachy RN, Fauquet C (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Molecular Biology 31: 1129-1139

War AR, Paulraj MG, Ahmad T, Buhroo AA, Hussain B, Ignacimuthu S, Sharma H C (2012) Mechanisms of plant defense against insect herbviores. Plant Signalling & Behaviour 7: 1306-1320

Wise I, Lamb R, Smith M (2001) Domestication of wheats (Gramineae) and their susceptibility to herbivory by Sitodiplosis mosellana (Diptera: Cecidomyiidae). The Canadian Entomologist 133: 255-67

Zheng S (1965) Wheat Midge. Beijing: Agricultural Press.

Claims

1. An isolated nucleic acid encoding for a protein conferring resistance to OWBM, wherein the nucleic acid comprises at least one of the groups consisting of:

    a) Group 1:

        i. A nucleic acid comprising SEQ ID NO:4 or 5, or
        ii. A nucleic acid encoding an amino acid sequence comprising SEQ ID NO:6, or
        iii. A nucleic acid encoding an amino-acid sequence comprising at least one CC motif, one NBSARC motif and one LRR motif and wherein the sequence encoding the LRR motif is having at least 82% identity with the nucleic acid fragment as depicted in SEQ ID NO:8, and

    b. Group 2:

        i. A nucleic acid comprising SEQ ID NO:1 or 2, or
        ii. A nucleic acid encoding an amino acid sequence comprising SEQ ID NO:3, or
        iii. A nucleic acid encoding an amino-acid sequence comprising at least one CC motif, one NBS-ARC motif and one LRR motif, and wherein the sequence encoding the LRR motif is having at least 78% identity with the nucleic acid fragment as depicted in SEQ ID NO:7,

2. An isolated amino acid sequence encoded by the nucleic acid of claim 1.

3. A nucleic acid construct comprising the nucleic acid of claim 1 wherein said nucleic acid is cloned downstream of a heterologous promoter functional in a plant cell.

4. A vector comprising the nucleic acid construct of claim 3.

5. A host cell containing at least the nucleic acid construct of claim 3.

6. A transgenic plant, or a part of a transgenic plant comprising at least one cell according to claim 5.

7. The transgenic plant of claim 6 wherein the plant is a wheat plant.

8. A method for conferring or improving resistance to OWBM in a plant, comprising the steps consisting of:

    a. transforming at least a plant cell or plant tissue with a vector containing, as a transgene, a nucleic acid

construct of claim 3;
b. cultivating the cell(s) or plant tissue thus transformed so as to generate a transgenic plant containing at least a cell which contains, in its genome, at least said nucleic acid construct.

wherein said transgenic plant presents a newly conferred or an improved resistance compared to a plant not comprising said nucleic acid construct.

9. A method for identifying a transgenic plant that can be used in a breeding process for obtaining a plant with a new or improved resistance to OWBM comprising the step of identifying, in a population of plants, the plants containing a nucleic acid construct of claim 3.

10. The method of anyone of claims 8 and 9, wherein said plant is a cereal and more preferentially is a wheat plant.

11. A marker within the Sm1 locus defined by Wa_c6957_32 and Excalibur c1787_1301 markers, wherein said Wa_c6957_32 marker is SEQ ID NO: 9 or SEQ ID NO: 10, and said Excalibur_c1787_1301 marker is SEQ ID NO: 11 or SEQ ID NO: 12, of the IWGSC Whole genome assembly ('IWGSC WGA') Chinese Spring genome and linked to Sm1 resistance.

12. A marker of claim 11 as listed in Table 1.

13. Method for identifying a plant resistant to OWBM compared to a plant lacking the nucleic acid of claim 1 wherein the method is comprising the steps of:

a. isolating DNA genomic sequence or RNA nucleic acid, and
b. identifying the presence of the nucleic acid of claim 1 or the presence of resistant alleles and/or the absence of susceptible alleles of a marker within the Sm1 locus defined by Wa_c6957_32 and Excalibur c1787_1301 markers, wherein said Wa_c6957_32 marker is SEQ ID NO: 9 or SEQ ID NO: 10, and said Excalibur c1787_1301 marker is SEQ ID NO: 11 or SEQ ID NO: 12, of the IWGSC Whole genome assembly ('IWGSC WGA') Chinese Spring genome.

14. The method of claim 13 wherein the resistant alleles are the sequences as listed in Table 1.

15. Use of SEQ ID NO: 7 or SEQ ID NO: 8 as a probe to identify and isolate orthologs of the RGA1 or RGA2 genes encoding respectively for SEQ ID NO:3 and SEQ ID NO:6 in a plant genome.

16. Use of the nucleic acid of claim 1 or a fragment of said nucleic acid as a screening tool for identifying a ligand peptide interacting with the protein encoded by nucleic acid of claim 1.

Patentansprüche

1. Isolierte Nukleinsäure, die für ein Protein codiert, das Resistenz für OWBM verleiht, wobei die Nukleinsäure mindestens eine der Gruppen umfasst, bestehend aus:

a) Gruppe 1:

i. eine Nukleinsäure, die SEQ ID Nr. 4 oder 5 umfasst, oder
ii. eine Nukleinsäure, die eine Aminosäuresequenz codiert, die SEQ ID Nr. 6 umfasst, oder
iii. eine Nukleinsäure, die eine Aminosäuresequenz codiert, die mindestens ein CC-Motiv, ein NBSARC-Motiv und ein LRR-Motiv umfasst, und wobei die Sequenz, die das LRR-Motiv codiert, mindestens 82% Identität mit dem wie in SEQ ID Nr. 8 dargestellten Nukleinsäurefragment aufweist; und

b) Gruppe 2:

i. eine Nukleinsäure, die SEQ ID Nr. 1 oder 2 umfasst, oder
ii. eine Nukleinsäure, die eine Aminosäuresequenz codiert, die SEQ ID Nr. 3 umfasst, oder
iii. eine Nukleinsäure, die eine Aminosäuresequenz codiert, die mindestens ein CC-Motiv, ein NBS-ARC-Motiv und ein LRR-Motiv umfasst, und wobei die Sequenz, die das LRR-Motiv codiert, mindestens 78%

Identität mit dem wie in SEQ ID Nr. 7 dargestellten Nukleinsäurefragment aufweist.

2. Isolierte Aminosäuresequenz, die durch die Nukleinsäure nach Anspruch 1 codiert wird.

3. Nukleinsäurekonstrukt, umfassend die Nukleinsäure nach Anspruch 1, wobei die Nukleinsäure stromabwärts von einem heterologen Promotor, der in einer Pflanzenzelle funktionell ist, geklont wird.

4. Vektor, umfassend das Nukleinsäurekonstrukt nach Anspruch 3.

5. Wirtszelle, enthaltend mindestens das Nukleinsäurekonstrukt nach Anspruch 3.

6. Transgene Pflanze oder Teil einer transgenen Pflanze, umfassend mindestens eine Zelle nach Anspruch 5.

7. Transgene Pflanze nach Anspruch 6, wobei die Pflanze eine Weizenpflanze ist.

8. Verfahren zum Verleihen oder Verbessern der Resistenz für OWBM in einer Pflanze, umfassend die Schritte bestehend aus:

a. Transformieren mindestens einer Pflanzenzelle oder eines Pflanzengewebes mit einem Vektor, enthaltend ein Nukleinsäurekonstrukt nach Anspruch 3 als ein Transgen;
b. Kultivieren der so transformierten Zelle oder Zellen oder des so transformierten Pflanzengewebes, um eine transgene Pflanze zu erzeugen, die mindestens eine Zelle enthält, die in ihrem Genom mindestens das Nukleinsäurekonstrukt enthält.

9. Verfahren zum Identifizieren einer transgenen Pflanze, die in einem Züchtungsprozess zum Erhalten einer Pflanze mit einer neuen oder verbesserten Resistenz für OWBM verwendet werden kann, umfassend den Schritt eines Identifizierens der Pflanzen, die ein Nukleinsäurekonstrukt nach Anspruch 3 enthalten, in einer Population von Pflanzen.

10. Verfahren nach einem der Ansprüche 8 und 9, wobei die Pflanze ein Getreide ist und mehr bevorzugt eine Weizenpflanze ist.

11. Marker innerhalb des Sm1-Locus, der durch einen Wa_c6957_32- und einen Excalibur_c1787_1301-Marker, wobei der Wa_c6957_32-Marker SEQ ID Nr. 9 oder SEQ ID Nr. 10 ist und der Excalibur_c1787_1301-Marker SEQ ID Nr. 11 oder SEQ ID Nr. 12 ist, der Chinese-Spring-Genom-IWGSC-Whole-Genome-Assembly ("IWGSC-WGA") definiert ist und mit Sm1-Resistenz verknüpft ist.

12. Marker nach Anspruch 11, wie in Tabelle 1 aufgeführt.

13. Verfahren zum Identifizieren einer Pflanze, die für OWBM resistent ist, im Vergleich zu einer Pflanze, der die Nukleinsäure nach Anspruch 1 fehlt, wobei das Verfahren die Schritte umfasst:

a. Isolieren einer genomischen DNA-Sequenz oder RNA-Nukleinsäure, und
b. Identifizieren des Vorliegens der Nukleinsäure nach Anspruch 1 oder des Vorliegens von resistenten Allelen und/oder des Fehlens von suszeptiblen Allelen eines Markers innerhalb des Sm1-Locus, der durch einen Wa_c6957_32- und einen Excalibur_c1787_1301-Marker, wobei der Wa_c6957_32-Marker SEQ ID Nr. 9 oder SEQ ID Nr. 10 ist und der Excalibur_c1787_1301-Marker SEQ ID Nr. 11 oder SEQ ID Nr. 12 ist, der Chinese-Spring-Genom-IWGSC-Whole-Genome-Assembly ("IWGSC-WGA") definiert ist.

14. Verfahren nach Anspruch 13, wobei die resistenten Allele die wie in Tabelle 1 aufgeführten Sequenzen sind.

15. Verwendung von SEQ ID Nr. 7 oder SEQ ID Nr. 8 als eine Sonde, um Orthologe des RGA1- oder des RGA2-Gens, das für SEQ ID Nr. 3 bzw. SEQ ID Nr. 6 codiert, in einem Pflanzengenom zu identifizieren und zu isolieren.

16. Verwendung der Nukleinsäure nach Anspruch 1 oder eines Fragments der Nukleinsäure als ein Screening-Werkzeug zum Identifizieren eines Ligandenpeptids, das mit dem Protein wechselwirkt, das durch eine Nukleinsäure nach Anspruch 1 codiert wird.

Revendications

1. Acide nucléique isolé codant une protéine conférant une résistance à la OWBM, lequel acide nucléique comprend au moins l'un des groupes consistant en :

    a) Groupe 1 :

        i. un acide nucléique comprenant la SEQ ID NO : 4 ou 5, ou
        ii. un acide nucléique codant une séquence d'acides aminés comprenant la SEQ ID NO : 6, ou
        iii. un acide nucléique codant une séquence d'acides aminés comprenant au moins un motif CC, un motif NBSARC et un motif LRR, et dans lequel la séquence codant le motif LRR a une identité d'au moins 82 % avec le fragment d'acide nucléique tel que représenté par la SEQ ID NO : 8,

    b) Groupe 2 :

        i. un acide nucléique comprenant la SEQ ID NO : 1 ou 2, ou
        ii. un acide nucléique codant une séquence d'acides aminés comprenant la SEQ ID NO : 3, ou
        iii. un acide nucléique codant une séquence d'acides aminés comprenant au moins un motif CC, un motif NBS-ARC et un motif LRR, et dans lequel la séquence codant le motif LRR a une identité d'au moins 78 % avec le fragment d'acide nucléique tel que représenté par la SEQ ID NO : 7.

2. Séquence d'acides aminés isolée codée par l'acide nucléique de la revendication 1.

3. Construction d'acide nucléique comprenant l'acide nucléique de la revendication 1, dans laquelle ledit acide nucléique est cloné en aval d'un promoteur hétérologue fonctionnel dans une cellule végétale.

4. Vecteur comprenant la construction d'acide nucléique de la revendication 3.

5. Cellule hôte contenant au moins la construction d'acide nucléique de la revendication 3.

6. Plante transgénique, ou partie d'une plante transgénique, comprenant au moins une cellule selon la revendication 5.

7. Plante transgénique selon la revendication 6, laquelle plante est un plant de blé.

8. Procédé pour conférer ou améliorer la résistance à la OWBM dans une plante, comprenant les étapes consistant à :

    a. transformer au moins une cellule végétale ou un tissu végétal avec un vecteur contenant, en tant que transgène, une construction d'acide nucléique selon la revendication 3 ;
    b. cultiver la ou les cellules ou le tissu végétal, ainsi transformés, de façon à générer une plante transgénique contenant au moins une cellule qui contient, dans son génome, au moins ladite construction d'acide nucléique,

    dans lequel ladite plante transgénique présente une résistance nouvellement conférée ou améliorée en comparaison avec une plante ne comprenant pas ladite construction d'acide nucléique.

9. Procédé pour identifier une plante transgénique pouvant être utilisée dans un processus de sélection pour obtenir une plante ayant une résistance nouvelle ou améliorée à la OWBM, comprenant l'étape d'identification, dans une population de plantes, des plantes contenant une construction d'acide nucléique selon la revendication 3.

10. Procédé selon l'une quelconque des revendications 8 et 9, dans lequel ladite plante est une céréale et plus préférentiellement est un plant de blé.

11. Marqueur à l'intérieur du locus Sm1 défini par les marqueurs Wa_c6957_32 et Excalibur_c1787_1301, dans lequel ledit marqueur Wa_c6957_32 est la SEQ ID NO : 9 ou la SEQ ID NO : 10, et ledit marqueur Excalibur_c1787_1301 est la SEQ ID NO : 11 ou la SEQ ID NO : 12, du génome Chinese Spring de l'ensemble de génomes entiers de l'IWGSC ("IWGSC WGA") et est lié à une résistance à Sm1.

12. Marqueur selon la revendication 11 tel que listé dans le Tableau 1.

13. Procédé pour identifier une plante résistante à la OWBM en comparaison avec une plante dépourvue de l'acide nucléique de la revendication 1, lequel procédé comprend les étapes suivantes :

a. isolement d'une séquence génomique d'ADN ou d'un acide nucléique de type ARN, et
b. identification de la présence de l'acide nucléique de la revendication 1 ou de la présence d'allèles résistants et/ou de l'absence d'allèles sensibles d'un marqueur à l'intérieur du locus Sm1 défini par les marqueurs Wa_c6957_32 et Excalibur_c1787_1301, dans lequel ledit marqueur Wa_c6957_32 est la SEQ ID NO : 9 ou la SEQ ID NO : 10, et ledit marqueur Excalibur_c1787_1301 est la SEQ ID NO : 11 ou la SEQ ID NO : 12, du génome Chinese Spring de l'ensemble de génomes entiers de l'IWGSC ("IWGSC WGA").

14. Procédé selon la revendication 13, dans lequel les allèles résistants sont les séquences telles que listées dans le Tableau 1.

15. Utilisation de la SEQ ID NO : 7 ou de la SEQ ID NO : 8 en tant que sonde pour identifier et isoler des orthologues des gènes RGA1 ou RGA2 codant respectivement la SEQ ID NO : 3 et la SEQ ID NO : 6 dans un génome de plante.

16. Utilisation de l'acide nucléique de la revendication 1 ou d'un fragment dudit acide nucléique en tant qu'outil de criblage pour identifier un peptide ligand interagissant avec la protéine codée par l'acide nucléique de la revendication 1.

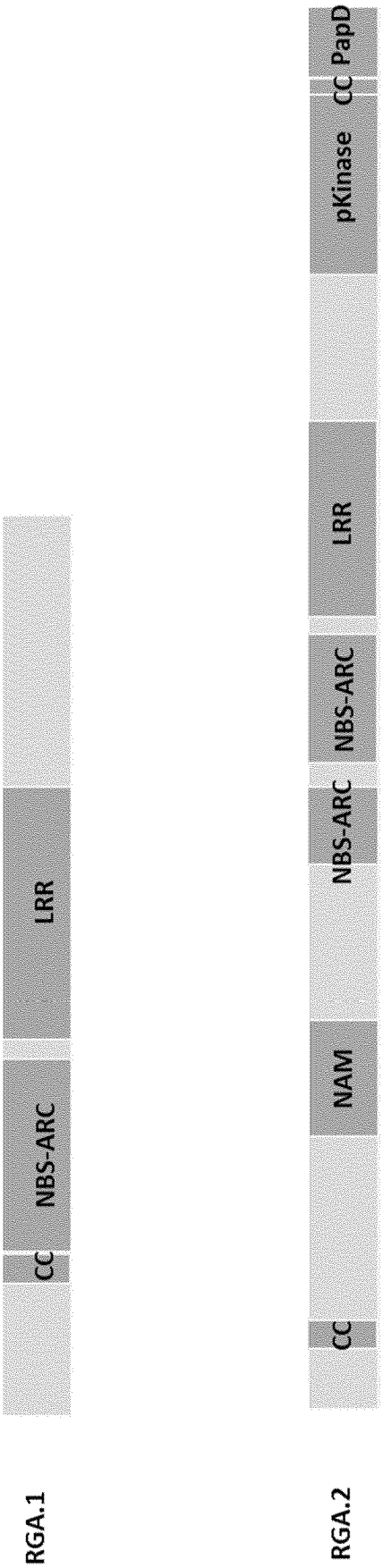
RGA.1
CC
NBS-ARC
LRR
RGA.2
CC
NAM
NBS-ARC
NBS-ARC
LRR
pKinase
CC
PapD

Figure 1

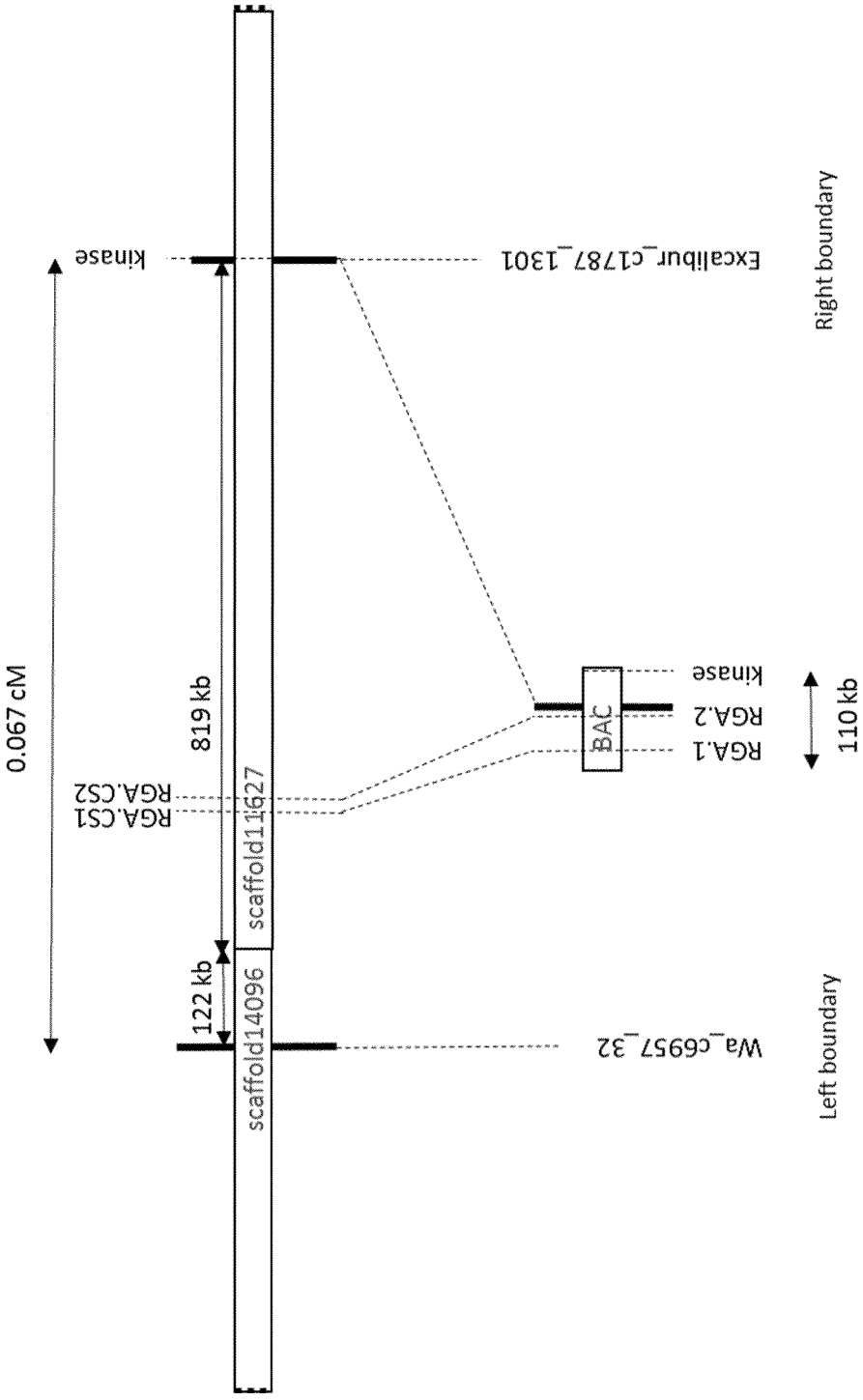
0.067 cM
819 kb
122 kb
scaffold14096
scaffold11627
RGA.CS2
RGA.CS1
kinase
Excalibur_c1787_1301
Right boundary
BAC
kinase
RGA.2
RGA.1
110 kb
Wa_c6957_32
Left boundary

Figure 2

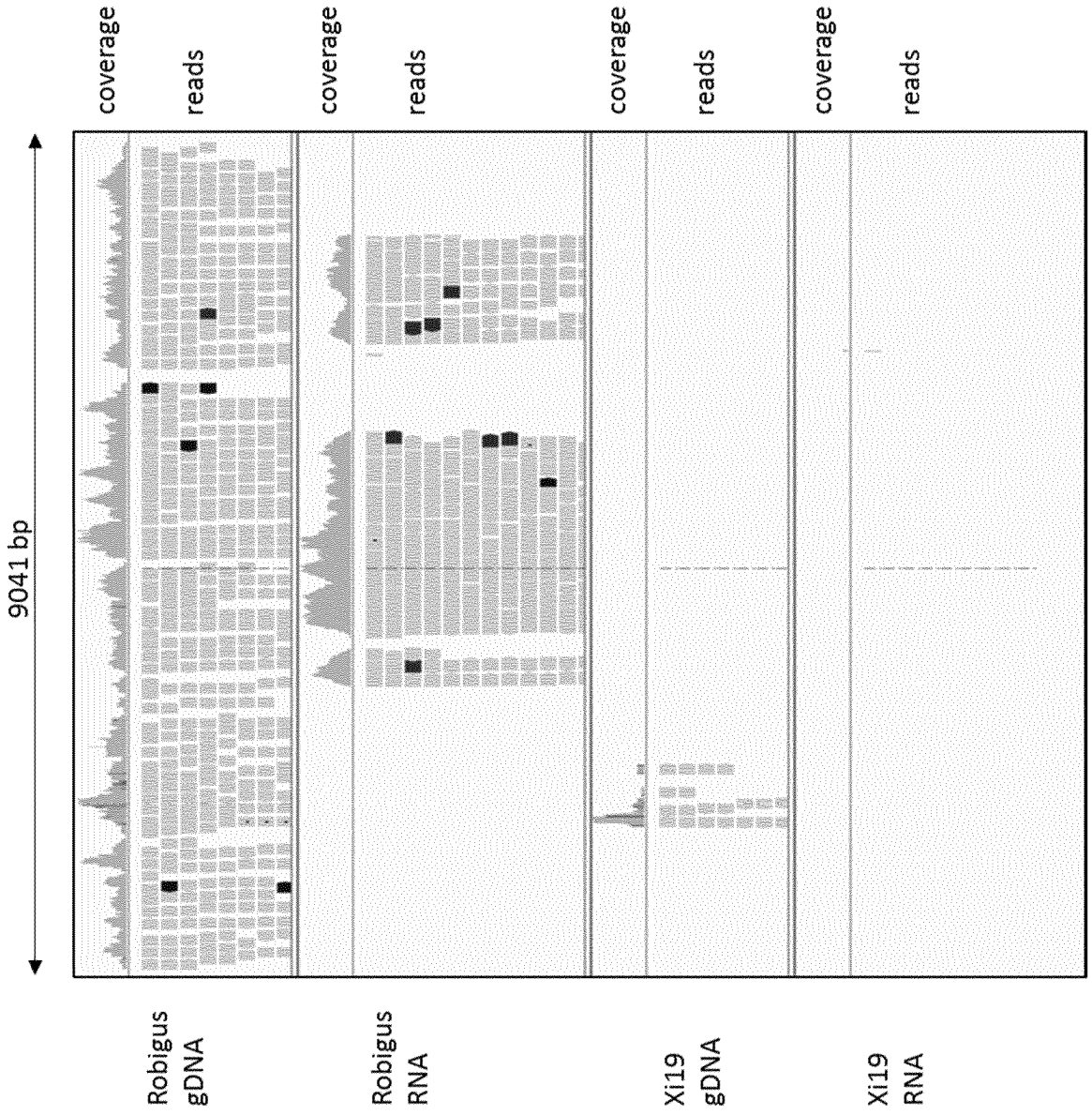
9041 bp
coverage
reads
coverage
reads
coverage
reads
coverage
reads
Robigus gDNA
Robigus RNA
Xi19 gDNA
Xi19 RNA

Figure 3

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 2011072246 A **[0033] [0117]**
- WO 2013181440 A **[0033] [0117]**
- US 20150007360 A **[0047]**
- US 20120011621 A **[0047]**
- US 20100306876 A **[0047]**
- US 20090307795 A **[0047]**
- US 20070028327 A **[0047]**
- US 20130024998 A **[0048]**

Non-patent literature cited in the description

- **ALTSCHUL SF ; MADDEN TL ; SCHAFFER AA ; ZHANG J ; ZHANG Z ; MILLER W ; LIPMAN DJ.** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.,* 1997, vol. 25, 3389-3402 **[0147]**
- **ALTSCHUL SF ; WOOTTON JC ; GERTZ EM ; AGARWALA R ; MORGULIS A ; SCHAFFER AA ; YU YK.** Protein database searches using compositionally adjusted substitution matrices. *The FEBS Journal,* 2005, vol. 272, 5101-5109 **[0147]**
- **ANDERSON OD ; GREENE FC.** The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of hexaploid bread wheat. *Theoretical and Applied Genetics,* 1989, vol. 77, 689-700 **[0147]**
- **ANON.** Orange wheat blossom midge. *AHDB Information sheet N° 53,* 2016, https://cereals.ahdb.org.uk **[0147]**
- **BAILEY TL ; ELKAN C.** Fitting a mixture model by expression maximization to discover in biopolymers. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology,* 1994, vol. 2, 28-36 **[0147]**
- **BIRKETT MA ; BRUCE TJA ; MARTIN JL ; SMART LE ; OAKLEY J ; WADHAMS LJ.** Responses of female orange wheat blossom midge to wheat panicle volatiles. *Journal of Chemical Ecology,* 2004, vol. 30, 1319-1328 **[0147]**
- **BLAKE NK ; STOUGAARD RN ; WEAVER DK ; SHERMAN JD ; LANNING SP ; NAROUKA Y ; XUE Q ; MARTIN JM ; TALBERT LE.** Identification of a quantitative trait locus for resistance to Sitodiplosis mosellana (Géhin), the orange wheat blossom midge, in spring wheat. *Plant Breeding,* 2011, vol. 130, 25-30 **[0147]**
- **BRUCE TJ ; HOOPER AM ; IRELAND L ; JONES OT ; MARTIN JL ; SMART LE ; OAKLEY J ; WADHAMS LJ.** Development of a pheromone trap monitoring system for orange wheat blossom midge, Sitodiplosis mosellana, in the UK. *Pest Management and Science,* 2007, vol. 63, 49-56 **[0147]**
- **CHAVALLE S ; CENSIER F ; SAN MARTIN ; GOMEZ G ; DE PROFTA M.** Protection of winter wheat against orange wheat blossom midge, Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae): efficacy of insecticides and cultivar resistance. *Pest Management Science,* 2015, vol. 71, 783-790 **[0147]**
- **CHRISTENSEN AH ; QUAIL PH.** Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research,* 1996, vol. 5, 213-218 **[0147]**
- **DEPIGNY-THIS D ; RAYNAL M ; ASPART L ; DELSENY M ; GRELLET F.** The cruciferin gene family in radish. *Plant Molecular Biology,* 1992, vol. 20, 467-479 **[0147]**
- **DING H ; LAMB RJ.** Oviposition and larval establishment of Sitodiplosis mosellana (Diptera: Cecidomyiidae) on wheat at different growth stages. *The Canadian Entomologist,* 1999, vol. 131, 475-481 **[0147]**
- **DING H ; LAMB RJ ; AMES N.** Inducible production of phenolic acids in wheat and antibiotic resistance to Sitodiplosis mosellana. *Journal of Chemical Ecology,* 2000, vol. 26, 969-985 **[0147]**
- Sitodiplosis mosellana (Géhin), orange wheat blossom midge (Diptera: Cecidomyiidae. **DOANE JF ; OLFERT OO ; ELLIOTT RH ; HARTLEY S ; MEERS S.** Biological control programmes in Canada 2001-2012. 2013, 272-276 **[0147]**
- **DOGIMONT C ; CHOVELON V ; PAUQUET J ; BOUALEM A ; BENDAHMANE A.** The Vat locus encodes for a CC-NBS-LRR protein that confers resistance to Aphis gossypii infestation and A. gossypii-mediated virus resistance. *The Plant Journal,* 2014, vol. 80, 993-1004 **[0147]**
- **DU B ; ZHANG W ; LIU B ; HU J ; WEI J ; SHI Z ; HE R ; ZHU L ; CHEN R ; HAN B.** Identification and characterization of Bph14, a gene conferring resistance to brown planthopper in rice. *Proceedings of the National Academy of Sciences USA,* 2009, vol. 106, 22163-22168 **[0147]**

- **ELLIOT RH ; MANN LW.** Susceptibility of red spring wheat, Triticum aestivum L. cv. Katepwa, during heading and anthesis to damage by wheat midge, Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae). *The Canadian Entomologist,* 1996, vol. 128, 367-375 **[0147]**
- **ELLIS SA ; BRUCE TJA ; SMART LE ; MARTIN JA ; SNAPE J ; SELF M.** Integrated management strategies for varieties tolerant and susceptible to wheat blossom midge. *HGCA project report No. 451,* 2009 **[0147]**
- **FIELDS S ; SONG O.** A novel genetic system to detect protein-protein interactions. *Nature,* 1989, vol. 340, 245-246 **[0147]**
- **GHARALARI AH ; FOX SL ; SMITH MAH ; LAMB RJ.** Ovipoistion deterrence in spring wheat, Triticum aestivum, against orange wheat blossom midge, Sitodiplosis mosellana: implications for inheritance of deterrence. *Entomologia Experimentalis et Applicata,* 2009, vol. 133, 74-83 **[0147]**
- **GHARALARI AH ; FOX SL ; SMITH MAH ; LAMB RJ.** The relationship between morphological traits of the spring wheat spike and oviposition deterrence to orange wheat blossom midge. *Entomologia Experimentalis et Applicata,* 2009, vol. 132, 182-190 **[0147]**
- **GHARALARI AH ; SMITH MAH ; FOX SL ; LAMB RJ.** Volatile compounds from non-preferred wheat spikes reduce oviposition by Sitodiplosis mosellana. *The Canadian Entomologist,* 2011, vol. 143, 388-391 **[0147]**
- **GRIES R ; GRIES G ; KHASKIN G ; KING S ; OLFERT OO ; KAMINSKI LA ; LAMB R ; BENNETT R.** Sex pheromone of orange wheat blossom midge, Sitodiplosis mosellana. *Naturwissenschaften,* 2000, vol. 87, 450-454 **[0147]**
- **HARRIS MO ; STUART JJ ; MOHAN M ; NAIR S ; LAMB RJ ; ROHFRITSCH O.** Grasses and gall midges: Plant defense and insect adaptation. *Annual Review of Entomology,* 2003, vol. 48, 549-577 **[0147]**
- **ISHIDA Y ; SAITO H ; OHTA S ; HIEI Y ; KOMARI T ; KUMASHIRO T.** High efficiency transformation of maize (Zea mays L.) mediated by Agrobacterium tumefaciens. *Nature Biotechnology,* 1996, vol. 14, 745-750 **[0147]**
- **JOUANIN L ; GUERCHE P ; PAMBOUKDJIAN N ; TOURNEUR C ; CASSE DELBART F ; TOURNEUR J.** Structure of T-DNA in plants regenerated from roots transformed with Agrobacterium rhizogene strain A4. *Molecular and General Genetics,* 1987, vol. 206, 387-392 **[0147]**
- **KASSA MT ; HAAS S ; SCHLIEPHAKE E ; LEWIS C ; YOU FM ; POZNIAK CJ ; KRAMER I ; PEROVIC D ; SHARPE AG ; FOBERT PR.** A saturated SNP linkage map for the orange blossom midge resistance gene Sm1. *Theoretical and Applied Genetics,* 2016, vol. 129, 1507-1517 **[0147]**
- **KAY R ; CHAN, A ; DALY M ; MCPHERSON J.** Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science,* 1987, vol. 236, 1299-1302 **[0147]**
- **LAMB RJ ; MCKENZIE RIH ; WISE IL ; BARKER PS ; SMITH MAH.** Resistance to Sitodiplosis mosellana (Diptera: Cecidomyiidae) in spring wheat (Gramineae). *The Canadian Entomologist,* 2000, vol. 132, 591-605 **[0147]**
- **LAMB RJ ; SMITH MAH ; WISE IL ; CLARKE P ; CLARKE J.** Oviposition deterrence to Sitodiplosis mosellana (Diptera: Cecidomyiidae): a source of resistance for durum wheat (Gramineae). *The Canadian Entomologist,* 2001, vol. 133, 579-591 **[0147]**
- **LAMB RJ ; WISE IL ; SMITH MAH ; MCKENZIE RIH ; THOMAS J ; OLFERT OO.** Oviposition deterrence against Sitodiplosis mosellana (Diptera: Cecidomyiidae) in spring wheat (Gramineae). *The Canadian Entomologist,* 2002, vol. 134, 85-96 **[0147]**
- **LAMB RJ ; SRIDHAR P ; SMITH MAH ; WISE IL.** Oviposition preference and offspring performance of a wheat midge Sitodiplosis mosellana (Géhin) (Diptera: Cecidomyiidae) on defended and less defended wheat plants. *Environmental Entomology,* 2003, vol. 32, 414-420 **[0147]**
- **LAMB RJ ; SMITH MAH ; WISE IL ; MCKENZIE RIH.** Resistance to wheat midge (Diptera: Cecidomyiidae) in winter wheat and the origins of resistance in spring wheat (Poaceae). *The Canadian Entomologist,* 2015, vol. 1, 1-10 **[0147]**
- **MCELROY D ; ZHANG W ; CAO J ; WU R.** Isolation of an efficient actin promoter for use in rice transformation. *The Plant Cell,* 1990, vol. 2, 163-171 **[0147]**
- **MCKENZIE RIH ; LAMB RJ ; AUNG T ; WISE IL ; BARKER P ; OLFERT OO.** Inheritance of resistance to wheat midge, Sitodiplosis mosellana, in spring wheat. *Plant Breeding,* 2002, vol. 121, 383-388 **[0147]**
- **MEYERS BC ; DICKERMAN AW ; MICHELMORE RW ; SIVARAMAKRISHNAN S ; SOBRAL BW ; YOUNG ND.** Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant Journal,* 1999, vol. 20, 317-332 **[0147]**
- **MILLER BS ; HALTON P.** The damage to wheat kernels caused by the wheat blossom midge (Sitodiplosis mosellana). *Journal of Science, Food & Agriculture,* 1960, vol. 12, 391-398 **[0147]**
- **MULLIS KB ; FALOONA FA.** Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology,* 1987, vol. 155, 335-350 **[0147]**
- **OAKLEY JN.** Orange wheat blossom midge: a literature review and survey of the 1993 outbreak. *Research Review No. 28, HGCA, Hamlyn House, Highgate Hill, London, UK,* 1994 **[0147]**

- **OAKLEY JN ; TALBOT G ; DYER C ; SELF MM ; FREER JBS ; ANGUS WJ ; BARRETT JM ; FEUERHELM G ; SNAPE J ; SAYERS L.** Integrated control of wheat blossom midge: variety choice, use of pheromone traps and treatment thresholds. *HGCA Project, Report 363,* 2005 **[0147]**
- **OLFERT O ; ELLIOTT RH ; HARTLEY S.** Non-native insects in agriculture: strategies to manage the economic and environmental impact of wheat midge, Sitodiplosis mosellana, in Saskatchewan. *Biological Invasions,* 2009, vol. 11, 127-133 **[0147]**
- **PERIYANNAN S ; MOORE J ; AYLIFFE M ; BANSAL U ; WANG X ; HUANG L ; DEAL K ; LUO M ; KONG X ; BARIANA H.** The gene Sr33, an ortholog of barley Mla genes, encodes the wheat stem rust race Ug99. *Science,* 2013 **[0147]**
- **ROBERT LS ; THOMPSON RD ; FLAVELL RB.** Tissue-specific expression of a wheat high molecular weight glutenin gene in transgenic tobacco. *The Plant Cell,* 1989, vol. 1, 569-578 **[0147]**
- **ROSSI M ; GOGGIN FL ; MILLIGAN SB ; KLAOSHIAN I ; ULLMAN DE ; WILLIAMSON VM.** The nematode resistance gene Mi of tomato confers resistance against the potato aphid. *Proceedings of the National Academy of Sciences USA,* 1998, vol. 95, 9750-9754 **[0147]**
- **SAMBROOK J ; FRITSCH EF ; MANIATIS T.** Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, 1989 **[0147]**
- **SARRIS PF ; CEVIK V ; DAGDAS G ; JONES JDG ; KRASILEVA KV.** Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biology,* 2016, vol. 14, 8 **[0147]**
- **SCHWEIGER W ; STEINER B ; VAUTRIN S ; NUSSBAUMER T ; SLEGWART G ; ZAMINI M ; JUNGREITHMEIER F ; GRATL V ; LEMMENS M ; MAYER KFX.** Suppressed recombination and unique candidate genes in the divergent haplotype encoding Fhb1, a major Fusarium head blight resistance locus in wheat. *Theoretical and Applied Genetics,* 2016, vol. 129, 1607-1623 **[0147]**
- **SEKHWAL KM ; PINGCHUAN L ; LAM I ; WANG X ; CLOUTIER S ; YOU FM.** Disease resistance gene analogs (RGAs) in Plants. *The International Journal of Molecular Sciences,* 2015, vol. 16, 19248-19290 **[0147]**
- **SMITH CM ; CLEMENT SL.** Molecular bases of plant resistance to arthropods. *Annual Review of Entomology,* 2012, vol. 57, 309-28 **[0147]**
- **STEMMER PC.** DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. *Proceedings of the National Academy of Sciences USA,* 1994, vol. 91, 10747-10751 **[0147]**
- High-throughput TILLING for functional genomics. **TILL BJ ; COLBERT T ; TOMPA R ; ENNS LC ; CODOMO CA ; JOHNSON JE ; REYNOLDS SH ; HENIKOFF JG ; GREENE EA ; STEIN MN.** Plant Functional Genomics: Methods and Protocols. Humana Press, 2003, 205-220 **[0147]**
- A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science,* 2014, vol. 345 (6194 **[0147]**
- **THOMAS J ; FINEBERG N ; PENNER G ; MCCARTNEY C ; AUNG T ; WISE I ; MCCALLUM B.** Chromosome location and markers of Sm1: a gene of wheat that conditions antibiotic resistance to orange wheat blossom midge. *Molecular Breeding,* 2005, vol. 15, 183-192 **[0147]**
- **VERDAQUER B ; DE KOCHKO A ; BEACHY RN ; FAUQUET C.** Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. *Plant Molecular Biology,* 1996, vol. 31, 1129-1139 **[0147]**
- **WAR AR ; PAULRAJ MG ; AHMAD T ; BUHROO AA ; HUSSAIN B ; IGNACIMUTHU S ; SHARMA H C.** Mechanisms of plant defense against insect herbviores. *Plant Signalling & Behaviour,* 2012, vol. 7, 1306-1320 **[0147]**
- **WISE I ; LAMB R ; SMITH M.** Domestication of wheats (Gramineae) and their susceptibility to herbivory by Sitodiplosis mosellana (Diptera: Cecidomyiidae. *The Canadian Entomologist,* 2001, vol. 133, 255-67 **[0147]**
- **ZHENG S.** Wheat Midge. Agricultural Press, 1965 **[0147]**